

*INTERNATIONAL JOURNAL OF
MODERN ENGINEERING
RESEARCH (IJMER)*

ISSN : 2249-6645



Volume 1 - Issue 1

Web : www.ijmer.com

Email : ijmer.editor@gmail.com

International Journal of Modern Engineering Research (IJMER)

Editorial Board

Executive Managing Editor

Prof. Shiv Kumar Sharma
India

Editorial Board Member

Dr. Jerry Van
Department of Mechanical, USA

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

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

Dr. Nouby Mahdy Ghazaly
Minia University, Egypt

Dr. Stanley John
Department of Textile Engineering, United Kingdom

Dr. Valfitaf Rasoul
Professor and HOD of Electromechanical, Russian

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

Dr. Manko dora
Associate professor of Computer Engineering, Poland

Dr. Ahmed Nabih Zaki Rashed
Menoufia University, Egypt

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

Associate Editor Member

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

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

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

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

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

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

Jasvinder Singh Sadana

M. TECH, USIT/GGSIPU, India

Dr. Bharat Raj Singh

Associate Director, SMS Institute of Technology, Lucknow

DR. RAVINDER RATHEE

C. R. P, Rohtak, Haryana

Dr. S. Rajendran

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

Mohd Abdul Ahad

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

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

Software Products Risk Assessment (SPRA) Tool for Determining Source Code Risks

Armen Keshishian¹, Hasan Rashidi²

¹Computer science Dept, Qazvin Azad University
Iran, Tehran

²Hasan Rashidi, Qazvin Azad University
Iran, Tehran

Abstract— Lack of the management of software risks is one of the main reasons of software project failure. In order to implement proper risk management processes, it is necessary to evaluate risks, based on the specified criteria. The process of the assessment of risks is a time-consuming process in software engineering. So Tools for automated the risks assessment is needed. In this paper, a method for automatic evaluation of some important measures of risk management is provided. Getting the physical address of the project and analyzing line by line is how this method works. In this analysis, the risks between classes and the internal risks of any class discovered using some criteria. These criteria that are used in this method are based on three-tier architecture. Finally, proposed method, provide some quantities which are represent the impact of the risks. For showing the efficiency of this method, a tool named SPRA is implemented. At the end of this paper, a comparison between two out puts is represented, one output is based on the manual method and the second one is the SPRA tools output. These results indicate that the proposed automation method can increase the accuracy of the assessment while it is optimizing the time and avoiding human errors.

Keywords— Software risks, risk management, risk assessment automation

I. INTRODUCTION

In recent years growing of the requirements in the industries, leads to increasing the complexity of the software and, in turn, it leads to amplification of the failure probability. As the proper software can guarantee the success of an industry, inefficient software can leads to the failure of the industry then. The reasons that are leading the failure of the software are called risk. Analyzing and assessment of risks can help to reduce the failure of the projects. [1] – [3]

The first step for analyzing the risks is the determination of the probability and the impact of the risk. The best way to determine the impact of risks is quantitative measurement, if the gathering of information from different resources was available. The most common way is calculating the expected monetary value. For the calculation of the EMV, the equation number 1 is used, where the EMV is stand for Expected Monetary Value. The Impact can also be calculated from the maximum of impact which is shown in equation 2 where P_i represented the probability of the maximum value. [4]

$$EMV = \text{Probability} \times \text{Impact} \quad (1)$$

$$\text{Impact} = \text{maxImpact} \times P_i \quad (2)$$

$$EMV = P_e \times P_i \times \text{maxImpact} \quad (3)$$

Given the equations above, it's possible to obtain the EMV using equation 3, where P_e represented the probability of an event. There is another parameter which is called Management Reserve or MR in short. MR is summation of relative EMV's for all anticipated threats and the equation 4 is used to calculate it. This parameter used to reduce the risks. [4]

$$MR = \sum (\text{probability}_i \times \text{impact}_i) \quad (4)$$

When threats identified and classified, the answer of the risks can be formulated and in that moment the Risk management plan will be complete. Figure 1 shows the risk management plan.

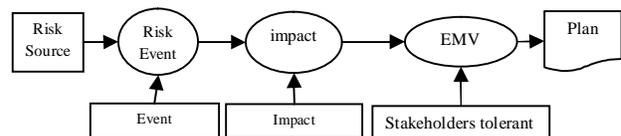


Fig. 1: Risk management plan [4]

Using some metrics for quantitative measurement is one of the most important methods in risk assessment. For that end, some important metrics that are lead to success or failure of a project should be gathered helping statistical information collection. Then for measuring the risks, the metrics should be calculated for any of the considered risks in the project. With comparison of the obtained values from the project with the values that are discovered in statistical method, the impact of any of the risks could be calculated. [4]

Now it is possible to use the impact parameter according to calculated impacts for each metric in equations 2, 3 and 4. Using these impacts and equations made it possible to give a value to each layer for assessment risks. With support of these quantitative measurements, the risk management plan will be so accurate and methodical. [4]

This paper introduces a new method that measures risk metrics in any software projects. At the end, the impacts of risks will be produced. For this goal, a tool that discovers all

of the risks has been developed. With using this tool, all of the impacts of the risks will be calculated.

There is different categorization for risks in various areas. In one of these categories that represented by Hoodat and Rashidi [7] risks have been divided into two different classifications that are Internal and external risks. Internal risks take place into an organization but externals occur out of it. For external risks some issues are proposed such as the market behavior, competitions, prices, widespread failure of the product and etc. Internal risks are classified into three different categories that are product, process and project. Project risks concern the performance of the project. The product risks include technical risks, pending faults and possible shortcomings. The process risks could be the outcome of product risks. It is important to determine the relations among the risks on these three categorize. [5]–[7]

In this paper one of the product risks is concerned that appears on the source code of the software. For avoiding some extra efforts, in the next section some similar works will be reviewed and then the new method will be introduced. At the end, the outputs of different models will be compared.

II. SIMILAR WORKS IN AUTOMATIC RISK ASSESSMENT

In this section some of the most important works that focus on discovering source codes risks will be reviewed. These efforts can be divided into two main approaches. These two approaches will be exhibited with their advantages and disadvantages.

A. DATRIX APPROACH

The name Datrix identifies a project, a set of tools and a team of engineers within Bell Canada. This approach tries to analysis the source code of the project. The aim of such an analysis is to assess the maintainability of these software products from a source code perspective. This model is based on the concept of an ASG, which stands for Abstract Semantic Graph. For generating this graph, the source code is parsed and the Abstract Syntax Tree (AST) is produced. The AST is then processed in order extract semantic information such as identifier scope, variable type and etc. This information is added to the AST as node attributes, edges or any other kind of annotations, that results in ASG production. In order to detect most risks, the ASG graph must be understandable. [8]

B. RISK ASSESSMENT USING SOURCE CODE APPROACH

For determining the risks, this approach divides the process into two phases. The information that is extracted in the first phase is generated using the automatic analysis of the source code. This information called primary information. Second phase results are called secondary information that are emerged from documents and the developers. A tool in Java language is developed to fetch the primary information using the source code analysis. The emerged information of the first phase then inserted into corresponding tables in database. In next stage, after the tool task, the analysis of the risks is on

analyzers to write appropriate queries to get the required information. [9]

Table 1 shows these two approaches advantages and disadvantages.

TABLE 1
comparison between two approaches [8],[9]

| Approach | Properties | Advantages | Disadvantages |
|-----------------------------------|---|---|--|
| Datrix | -using ASG -generating a graphical structure | -reducing reviewing process -representing a tree schema -the first graphical approach | -lack of the global view - lack of the version determining tool -lack of representing the sub graphs -tight dependency between the results and the analyzers -lack of some metrics for risk assessment |
| The source code analysis approach | -using database for storing the data | -reducing reviewing process -supporting the querying on the analyzed data | -high dependency between the results and the documents -high dependency between the results and the analyzers -lack of some metrics for risk assessment |

III. DISCOVERING PROJECT RISKS METHOD

This paper represents a new method which uses flexible parsing [10] and source code analyzer [11]. After the analyzing phase, the risky patterns are discovered using some methods in resources [11, 12]. Then the impacts of the projects risks would be calculated. At the end of the method, with using resource [13] some documents will be generated.

The source code risks are categorized in 2 different classes:

- Intercommunicated class risks
- Single class risks

A. INTERCOMMUNICATED CLASS RISKS

For assessing the risks in a project it is necessary to discover the communicated classes and exhibit the interaction with numbers. Discovering and assigning these relations could be helpful in predicting the propagation of the changes among the classes of the project. These values can be used to determine the percentage of the changes, according to the variations percentage of the classes. This may be useful in decision support system (DSS). The accessibility of these percentages can assist managers to decide whether to accept or reject the changes. In case of acceptance, the classes that will change are determined. Two different kinds of dependencies are considered for analyzing the changes:

- Hard Dependencies
- Soft Dependencies

In this article the hard dependencies are the inheritances between two classes. These kinds of dependencies are more important than the soft dependencies, because the most tightly

relation is created between these two classes when a class inherits from the other class. Therefore, the hard dependency value is higher than the soft dependency. The dependency value shows the propagation changes domain. In other words, it shows the number of affected classes. For discovering the hard dependencies in this article, the parent class is found using code analyzing method and then an inheritance relation is established among parent and the child classes.

The other kind of dependencies is soft dependency which contains the call of other class methods and usage of public members of the other classes. In this case a soft relation is established between two dependent classes. For this aim, first the public members of all classes are discovered and a list of those members is generated. After that, all of the classes will be analyzed with using the source code reviewing method. If any of items of the list is found in the class body, then a soft relation between two classes (*member owner class and user class*) will be created.

Discovering both dependencies has $O(n*m)$ time complexity. First of all the whole codes are parsed, and the public members are fetched from the source code and a list is generated, this step has $O(n)$ time complexity where n is the number of total code lines. At the second step, the code will be reviewed again and the generated list will be parsed simultaneously for each line. So the total time complexity will be $O(n*m)$ where n is the number of code lines and m is the summation of the public members and the number of all classes.

B. SINGLE CLASS RISKS

As the inefficiency of a single class can threaten the whole project, it is important to assess a class without considering other classes and their relations. For automatic assessment of this kind of risks, some of metrics have to be prepared. These metrics have been gathered from previous articles. These metrics are identified by some keywords. With linear parsing and comparing each statement with these keywords is the routine of this method. In this analyzing the usage percentage of the keywords will be obtained. With these percents and using some boundary values that are represented in other articles, values which are the impacts of all risks will be calculated.

The metrics that are used in this paper are derived from resources [15], [16] and [17] where each one has its own boundary values. These boundary values are obtained from the statistically gathering data method. Every metric has its own value which represents the risk impact. Comparing the calculated value of each metric with the corresponding statistical values could lead to new value retrieval which represents the risk impact of that metric. Table 2 shows the metrics and the boundary values and the scale of each one in ten.

TABLE 2
some important metrics and the boundary values

| # | Metric Name | Boundary Value | Ten Scale |
|---|---------------------------------------|----------------|-----------|
| 1 | Exception Handling Structures | Less than 5% | 8 |
| | | 5% to 8% | 7 |
| | | 8% to 12% | 5 |
| | | More than 12% | 1 |
| 2 | The percent of using comments | Less than 10% | 6 |
| | | 10% to 15% | 5 |
| | | More than 15% | 1 |
| 3 | The percent of using global variables | Less than 10% | 5 |
| | | 10% to 20% | 6 |
| | | 20% to 40% | 7 |
| | | More than 40% | 10 |
| 4 | Number of methods | Less than 20 | 1 |
| | | 20 to 40 | 3 |
| | | More than 40 | 5 |
| 5 | The percent of using standard objects | Less than 20% | 5 |
| | | 20% to 40% | 3 |
| | | 40% to 60% | 5 |
| | | More than 60% | 3 |

IV. SOFTWARE PRODUCTS RISK ANALYSER TOOL

Software products risk analyzer or SPRA in short, is an application that is developed to automatically discover risks. The main goal of this tool is to avoid the analyzers from reviewing massive source codes. This tool identifies the whole risky patterns with given source codes then a document is created with results of this identification. This document includes all of the risky metrics of the three-tier architecture projects with their impact values. The output of this application can be generated in a short time without human errors.

The SPRA is represented, according to section 2 and the weaknesses of mentioned methods. This tool is developed by C# programming language. With obtaining the physical address of the project as an input, the parser starts the analyzing with proposed method. The parser analyzes the whole code in linear manner.

The parser creates a class object for any file that contains a class. This object is a defined class and its properties are specified based on the common characteristics between classes. For example all of classes can inherit and they have some methods, and etc. Fig. 2 shows the schema of class object and its interaction with parent class.

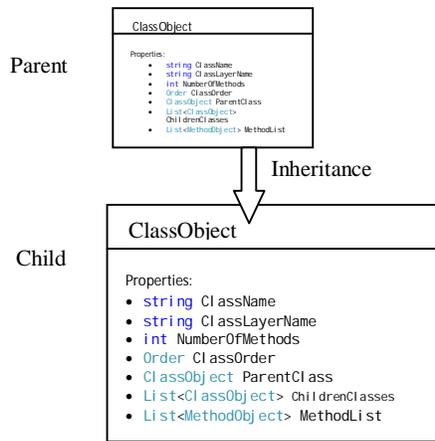


Fig. 2: the Risk Object and interaction with parent class

In this section, the risks of the layers are discussed. The purpose of using three-tier architecture is to apply the advantages of modular software development. In this architecture, any layer has its own duties; thus, there are some risks which threaten the related layer and the threat doesn't scatter to the other layers. So it is much easier to concentrate only on a layer risks instead of monitoring all risks of all layers. The SPRA tool applies this advantage and some part of it only deals with the risks of a layer.

A. Hard dependencies

There are lots of classes for any software which is developed by Object Oriented model. These classes usually have tight interconnections. One of the interconnections is inheritance relation that can cause some risks. Requirement changes or malfunctioning of a parent class could lead to the propagation of changes among wide variety of classes. For this reason, it is so vital to discover and monitor this kind of connection among classes. Figure 3 shows the inheritance between some classes and the propagation of changes in that set.

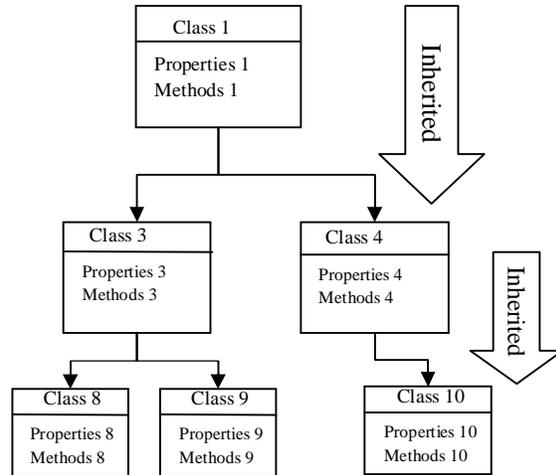


Fig. 3: the propagation of changes in inherited classes

At the beginning of class analysis, the process identifies all classes is determined. For this property a field with integer type has been considered. Value zero exhibits the protected type and value one shows the public type of the class modifier. With using this value, the scale of the class could be determined. The parent of classes is determined in parallel manner while the parser checks the modifier type, and if the parent class is one of the internal classes, then two classes will be connected to each other. This connection would be as an inherited type. There is a certain property for class object which is the parent class. This property will be initialized with the parent class object. Using this information could help to the prediction of the required changes in other classes. Table 3 shows the required fields for discovering the hard dependencies.

TABLE 3
required fields for discovering the hard dependencies

| # | Type | Field Name |
|---|-------------------|--------------|
| 1 | Int | Modifier |
| 2 | ClassObject | ParentClass |
| 3 | List<ClassObject> | ChildrenList |

B. Soft dependencies

Inheritance relations are not the only way in associating the classes. There is another kind of dependencies which is called soft dependency. For tracking this kind of connections, it has to have some certain fields in class object. To this end the UsedClasses field is considered as a list of class objects. For tracking the usage of other class methods, another field is considered as OtherClassesMethods. For used members, OtherClassesVars field is intended. With using these fields, the SPRA tool could track the propagation of changes of any class through the connected classes. This has $O(n)$ time complexity. Table 4 shows the related fields in class object.

TABLE 4
related fields in class object

| # | Type | Field Name |
|---|--------------------|---------------------|
| 1 | List<ClassObject> | UsedClasses |
| 2 | List<MethodObject> | OtherClassesMethods |
| 3 | List<Variable> | OtherClassesVars |

V. SPRA COMPONENTS

For a better understanding of SPAR operation, it is needed to review its components. The core of this tool is based on three classes. In the next section these classes will be presented in details, after that a comparison between the SPRA and the manual outputs will be made.

A. BOUNDARY CLASS

According to the similar researches, the lack of the risk assessment is one of the main inadequacies. With using the measured criteria, risk assessments can be reduce the duplicated processes and even can help analyzers to manage the risks in a better way. This facility is considered in SPRA tool. SPRA uses the output of other researches in which the data is gathered from very different projects. After the comparison and value selection for any metric, the SPRA tool creates a document and reports all of the risky parameters. In SPRA tool a class is considered for maintaining the boundary values. This class is a data structure that contains all of the keywords and corresponding values. Having this class and other components used in SPRA, makes it possible to obtain the risks of any arbitrary set of class. Table 5 shows the boundary values and some of the most important fields of this class.

TABLE 5
The boundary values and some of the most important fields of this class

| Field Name | Field Type | Values |
|-----------------|------------|--------|
| ExceptionUnder5 | int | 8 |
| Exception5to8 | int | 7 |
| ExceptionUpto12 | int | 1 |
| CommentUnder10 | int | 6 |
| Comment10To15 | int | 5 |
| NOMUnder20 | int | 1 |
| NOM20to40 | int | 3 |

B. AVERAGE CLASS

Average class is a static class which is used for determining the average of the risks of any set of classes. This arbitrary set could be a single class or all of certain layer classes or any other combination of classes. With helping the boundary value class, the average class calculates an average value of the given classes risks impacts.

There are several methods defined in this class to make it easy to get all type of risks. For example a method is developed for determining the risks related to exception handling. With helping the boundary values and using the corresponding class, the impacts can be calculated. The

average class calculates an average of these values and the result is stored in Risk Class.

C. RISK CLASS

For storing the assessed values, the risk class is introduced. This class has several fields for storing the values that will be used by managers and risk analyzers. To this end, this class uses three other subclasses. These classes are ClassMetrics, ProjectMetrics and LayerMetrics. In ClassMetrics some fields are considered for storing the assessed risks of any class object. The ProjectMetrics Class has the appropriate fields for maintaining the impacts of project level risks. The LayerMetrics class has some other fields for storing the impacts of any Layer risks. Figure 4 shows this class fields.

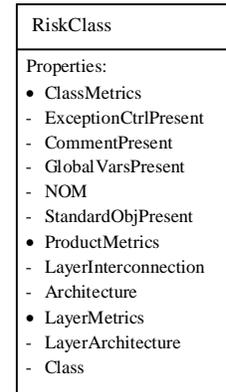


Fig. 4: Risk Class Fields

D. THE COMPONENTS INTERACTION

For better understanding to how this method works, the pseudo code 1 is represented. As pseudo code 1, at the first phase of the tools execution all of the required classes are added to AllClassesList object. After initializing this object a Risk Class object is created. Then for each item on AllClassesList the criteria of risks is calculated and the corresponding field is initialized. This initialization is handled by the Average class. The Average Class applies the boundary value class. Furthermore, using boundary values for each metric, the impact can be determined and stored in the corresponding subclass. Finally, using the documenting methods, a document is stored in the physical address of the project. This document contains all of the information that can be usefully in risk assessment and analyzing.

```

List<ClassObject> AllClassesList = initializing the favorite set of classes;
RiskClass riskClass = new RiskClass();
foreach (ClassObject cls in AllClassesList)
{
    riskClass.ClassMetric.ExceptionCtrlPresent =
        Average.ExceptionCtrlPresent(AllClassesList);
    riskClass.ClassMetric.CommentPresent =
        Average.CommentPresent(AllClassesList);
    riskClass.ClassMetric.GlobalVarPresent =
        Average.GlobalVarPresent(AllClassesList);
    riskClass.ClassMetric.NOM = Average.NOM(AllClassesList);
    riskClass.ClassMetric.StandardObjPresent =
        Average.StandardObjPresent(AllClassesList);
    riskClass.ProductMetric.LayerInterconnection =
        Average.LayerInterConnection(AllClassList);
    riskClass.LayerMetrics.LayerArchitecture =
        Average.LayerArchitecture(AllClassList);
    riskClass.LayerMetrics.Class = Average.Class(AllClassList);
}
    
```

Pseudo code 1: the interaction between components

VI. COMPARISON BETWEEN SPRA AND OTHER METHODS

For measuring the performance of the SPRA tool, some sample projects of Kaloob Engineering Corporation is selected. The Kaloob Corporation is a software development company that develops GIS based systems. For comparison, three different methods are considered. First method is the SPRA method output and the other one is based on the old methods that were mentioned, and the last one is the manual method that is implemented by the developers and managers. The old methods are obtained from references [11] and [12]. The SPRA metrics and boundary values are obtained from reference [10] and the manual metrics are utilized by managers and the risk analyzers. These comparisons are made using five different projects. Projects number 1, 2 and 3 are web-based and the others are windows-based. The best method is the one that its results are closer to the manual outputs. Figure 1 shows the result.

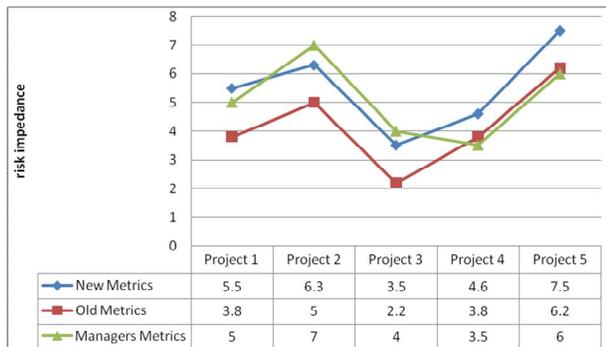


Fig. 5: some different methods outputs

According to this Diagram, the outputs of the SPRA tool are accurate enough to be used in real environments. This means that the SPRA could avoid the human errors while it

saves the assessment time. It can also be replaced with the manual assessing methods.

VII. CONCLUSION

In this article a method is introduced that can estimate the risks of software projects. The SPRA tool is developed based on the represented method to automate the risk assessments processes. The outputs show, SPRA could be replaced with the time-consuming manual methods. The output diagram shows that the type of the project could play an important role in risk assessment; thus, gathering the boundary values of impacts based on the project type could be one of the future researches.

REFERENCES

- [1] B. Xinlong, "Software Engineering Failures: A Survey". Oregon State: University Corvallis, School of EECS, 2001.
- [2] B. Lawhorn. (2009) Software Project Failure Costs Billions. Better Estimation & Planning Can Help. [Online]. <http://www.standish.com>.
- [3] W. Humphrey, "Five reasons why software projects fail". Addison-Wesley, 2002.
- [4] J. Kontio, "The riskit method for software risk management", Institute for Advanced Computer Studies and Department of Computer science, University of Maryland, 1999.
- [5] R. Pressman, *Software Engineering: A Practitioner's Approach*. 7th ed., McGraw-Hill Science Engineering, 2009.
- [6] H. Ronald, P. Haimés and Y. Yacov, "Software Risk Management". University of Virginia: Software Engineering Institute, Center for Risk Management of Engineering, 1996.
- [7] H. Hoodat, H. Rashidi. "Classification and Analysis of Risks in Software Engineering". World Academy of Science, Engineering and Technology 56, 2009.
- [8] Lapiere, Sebastien, Lague, Bruno and Leduc, Charles. "Datrix Source Code Model and its Interchange Format: Lessons Learned and Considerations for Future Work. Montreal", Canada: Bell Canada, Quality Engineering and Research, 2002.
- [9] van Deursen, Arie and Kuipers, Tobias. "Source-Based Software Risk Assessment". Netherlands: CWI and Delft University of Technology the Netherlands, 2004.
- [10] G. Knapen, B. Laguë, M. Dagenais, E. Merlo, "Parsing C++ Despite Missing Declarations", International Workshop on Program Comprehension, May 99, Pittsburgh, PA, USA.
- [11] M. Kuhnemann, T. Rauber, G. Runger, "A source code analyzer for performance prediction", Parallel and Distributed Processing Symposium, 18th International, 2004.
- [12] Li, Z. Lu, S. Myagmar, S. Zhou, "finding copy-paste and related bugs in large-scale software code", Software Engineering, IEEE Transactions on, March 2006.
- [13] Victor R. Basili and Salwa K. Abd-El-Hafiz, "A method for documenting code components", Journal of Systems and Software, Volume 34, Issue 2, August 1996, Pages 89-104.
- [14] H. Koziolka and F. Brosch, "Parameter Dependencies for Component Reliability Specifications", Sixth International Workshop on Formal Engineering approaches to Software Components and Architectures, FESCA, 2009.
- [15] A. Keshishian, H. Rashidi, "Assessing several Metrics for risk management in software", Sixteenth conference of Computer science, Sharif University of Tehran., 2010.
- [16] A. Stale, "Risk-based testing: Risk analysis fundamentals and metrics for software testing including a financial application case study", Systems and Software, pp. 287-295. 2000.
- [17] L. Rosenberg, H. Stapko and G. Albert. "Risk-based Object Oriented Testing", SATC NASA, Unisys, 2000.

Design and Implementation of Viterbi Decoder Using FPGA

Shoban Mude¹, S Nagakishore Bhavanam²

¹ Guru Nanak Institute of Technology, E.C.E Dept., Hyderabad, India

Email: shoban.mude@gmail.com

² Guru Nanak Institute of Technology, Asst. Prof, E.C.E Dept., Hyderabad, India

Email: kishorerreddy.vlsi@gmail.com

ABSTRACT

Convolutional encoding is a forward error correction technique that is used for correction of errors at the receiver end. The two decoding algorithms used for decoding the convolutional codes are Viterbi algorithm and Sequential algorithm. Sequential decoding has advantage that it can perform very well with long constraint length. Viterbi decoding is the best technique for decoding the convolutional codes but it is limited to smaller constraint lengths.

It has been widely deployed in many wireless communication systems to improve the limited capacity of the communication channels. The Viterbi algorithm is the most extensively employed decoding algorithm for convolutional codes. The availability of wireless technology has revolutionized the way communication is done in our world today. With this increased availability comes increased dependence on the underlying systems to transmit information both quickly and accurately. Because the communications channels in wireless systems can be much more hostile than in "wired" systems, voice and data must use forward error correction coding to reduce the probability of channel effects corrupting the information being transmitted. A new type of coding, called Viterbi coding, can achieve a level of performance that comes closer to theoretical bounds than more conventional coding systems. The Viterbi Algorithm, an application of dynamic programming, is widely used for estimation and detection problems in digital communications and signal processing. It is used to detect signals in communications channels with memory, and to decode sequential error control codes that are used to enhance the performance of digital communication systems.

Though various platforms can be used for realizing Viterbi Decoder including Field Programmable Gate Array (FPGAs), Complex Programmable Logic Devices (CPLDs) or Digital Signal Processing (DSP) chips but in this project benefit of using an FPGA to Implement Viterbi Decoding Algorithm has been described. FPGAs are a technology that gives the designer flexibility of a programmable solution, the performance of a custom solution and lowering overall cost. The advantages of the FPGA approach to DSP Implementation include higher sampling rates than are available from traditional DSP chips, lower costs than an ASIC. The FPGA also adds design flexibility and adaptability with optimal device utilization conserving both board space and system power that is often not the case with DSP chips.

Keywords: Convolutional encoding, Viterbi decoder, Path Metric, branch metric, FPGA, Xilinx, Modelsim.

I. INTRODUCTION

In digital communication system, error detection and error correction is important for reliable communication. Error detection techniques are much simpler than forward error correction (FEC). But error detection techniques have certain disadvantages. Error detection presupposes the existence of an automatic repeat request (ARQ) feature which provides for the retransmission of those blocks, segments or packets in which errors have been detected.

This assumes some protocol for reserving time for the retransmission of such erroneous blocks and for reinserting the corrected version in proper sequence. It also assumes sufficient overall delay and corresponding buffering that will permit such reinsertion. The latter becomes particularly difficult in synchronous satellite communication where the transmission delay in each direction is already a quarter second. A further drawback of error detection with ARQ is its inefficiency at or near the system noise threshold. For, as the error rate approaches the packet length, the majority of blocks will contain detected errors and hence require retransmission, even several times, reducing the throughput drastically. In such cases, for forward error correction, in addition to error detection with ARQ, may considerably improve throughput. Forward error correction may be desirable in place of, or in addition to, error detection for any of the following reasons:

- I. When a reverse channel is not available or the delay with ARQ would be excessive.
- II. There transmission strategy is not conveniently implemented.

Keeping in view requirements of communication channels in 3G wireless systems, need of reliable data communication, fast as well as accurate is the main consideration. So main work in this Paper is to develop an algorithm for a Viterbi Decoder to decode an encoded data.

II. BACKGROUND

Error correction coding [9] can be used to detect and correct data transmission errors in communication channels. Encoding is accomplished through the addition of redundant bits to transmitted information symbols.

These redundant bits provide decoders with the capability to correct transmission errors. Convolutional codes [7] form a set of popular error-correction codes. In convolutional coding, the encoded output of a transmitter (encoder) depends not only on the set of encoder inputs received during a particular time step, but also on the set of inputs received within a previous span of $K-1$ time units, where K is greater than 1. The parameter K is the *constraint length* of the code.

A typical convolutional encoder of constraint length $K = 3$ is shown in Figure 1. As shown in the figure, the encoding of convolutional codes can be accomplished with shift registers and generator polynomials (XOR functions). A convolutional encoder is represented by the number of output bits per input bit (v), the the number of input bits accepted at a time (b), and the constraint length (K), leading to representation (v, b, K) . Figure 1 depicts a $(2, 1, 3)$ convolutional encoder since the encoder accepts one input bit per time step and generates two output bits. The two output bits are dependent on the present input and the previous two input bits. The constraint length K indicates the number of times each input bit has an effect on producing output bits. Larger constraint lengths, i.e. $K = 9$ or higher, are preferable since they allow for more accurate error correction. Encoding *rate* R is equal to b/v . In many communication systems, a rate of $1/2$ is used [9]. Initially, the contents of the encoder shift register are set to zero. The contents are shifted right each time a one-bit value is converted into a two-bit symbol and transmitted. The operation of the encoder can be represented by a *state diagram*, as shown in Figure 2. Nodes represent the present state of the shift register while edges represent the output sequence and point to the next state of transition. Successive evaluation of state over time leads to the trellis diagram shown in Figure 3.

The diagram is a time-ordered mapping of encoder state with each possible state represented by a point on the vertical axis. Nodes represent the present state of the shift register at specific points in time while edges represent the output sequence and point to the next state of transition. The horizontal axis represents time steps. Branch lines indicate the transition of the present state of the shift register to the next state upon receiving a particular input bit, b . The upper branch leaving a node implies an input of 0 while the lower branch implies an input of 1.

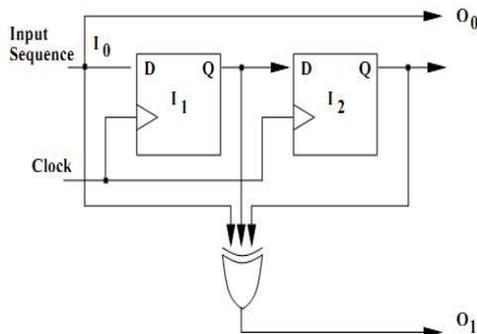


Fig.1 A(2,1,3) Convolution Encoder

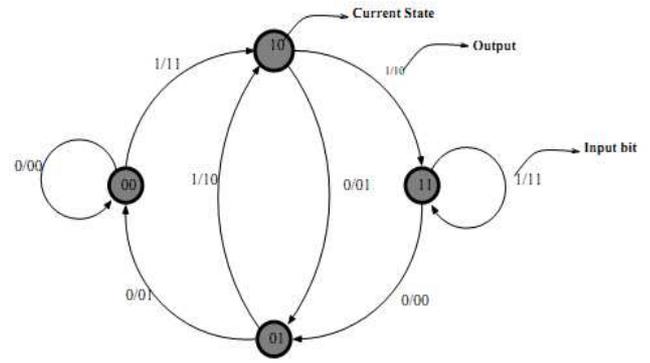


Fig.2 State Diagram for the Convolutional Encoder in Fig.1

The function of the decoder is to attempt to reconstruct the input sequence transmitted by the encoder by evaluating the received channel output. Values received at the decoder may differ from values sent by the encoder due to channel noise. The interaction between states represented by the trellis diagram is used by a decoder to determine the likely transmitted data sequence [16] as v -bit symbols are received. An example received sequence of two-bit v values appears at the top of Figure 3. The cost of a particular transition edge (branch) is determined from the Hamming distance of the received symbol and the *expected* symbol, labelled in bold on the transition edge. At each node the cumulative cost or *path metric* of the path is determined. These values are labeled in bold at each node in the figure. If multiple paths converge on the same state, the lowest cost path is preserved and other paths are eliminated. After a series of time steps, referred to as the *truncation length* (TL), the lowest-cost path, also known as minimum distance path, is determined, identifying the most-likely transmitted symbol sequence. The typical value of the truncation length depends on the noise in the channel and has been empirically found to be 3-5 times the constraint length [9]. Each *path* in the trellis diagram represents a unique set of inputs, such as the path highlighted in bold, corresponding to the lowest-cost input sequence $b = (0110)$.

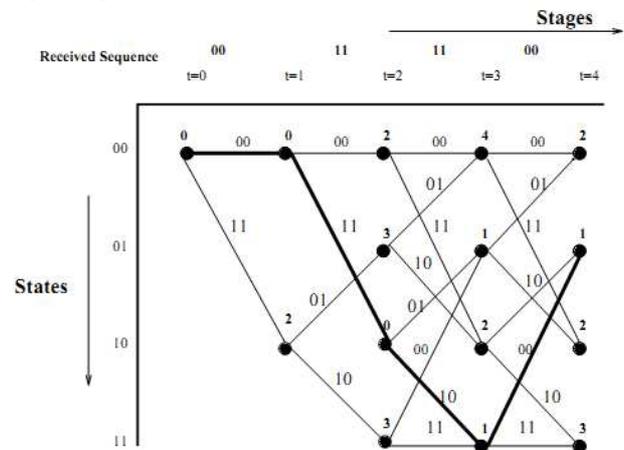


Fig.3 Trellis Diagram for the Convolutional Encoder in Fig.1

which satisfy the threshold condition may be discarded, potentially leading to a large BER. Alternately, if N_{max} is set to a large value, extra computation and memory are required, potentially with little benefit to BER reduction. As a result, an optimal value for N_{max} should be chosen to balance hardware size and BER.

Several reduced-complexity algorithms similar to the adaptive Viterbi algorithm have been developed, although each has significant limitations. The M-algorithm [6] is a popular reduced-complexity alternative to the Viterbi algorithm. Like the AVA, complexity reduction is achieved by retaining only the best M (N_{max}) paths at each trellis stage. Unlike AVA, this approach does not use a threshold condition to determine which paths are saved but rather sorts all paths and retains the M lowest-cost paths. This requirement of sorting circuitry adds complexity and delay to the M-algorithm. Since the AVA only requires comparison to a value determined by the metric of the best extended path, sorting is not required. A *beam-search* algorithm, which is similar to the AVA, was implemented in software in [10]. This approach was used in conjunction with hidden Markov modelling (HMM) of speech.

Unlike previous AVA approaches [4], the standard operation of eliminating the largest-metric path when two survivor paths enter the same trellis state was not implemented in our approach due to hardware

complexity. The implemented algorithm more closely resembles a variant of the AVA known as the Simmons T-algorithm [11].

IV. AVA ARCHITECTURE

To demonstrate the benefit of the adaptive Viterbi algorithm we have developed the first hardware implementation of the algorithm. This architecture takes advantage of parallelization and specialization of hardware for specific constraint lengths and dynamic reconfiguration to adapt decoder hardware to changing channel noise characteristics.

Description of the architecture:

The architecture of the implemented adaptive Viterbi decoder is shown in Figure 5 for the encoder with parameters shown in Figure 1, (2, 1, 3). The adaptive Viterbi decoder accepts two inputs from the channel which represent the outputs of the encoder that have been transmitted. The *branch metric generator* determines the difference between the received v -bit (in this case 2) value and $2v$ possible expected values. This difference is the Hamming distance between the values. A total of $2v$ branch metrics are determined by the branch metric generator. For $v=2$ these metrics are labelled b_{00}, b_{01}, b_{10} and b_{11} .

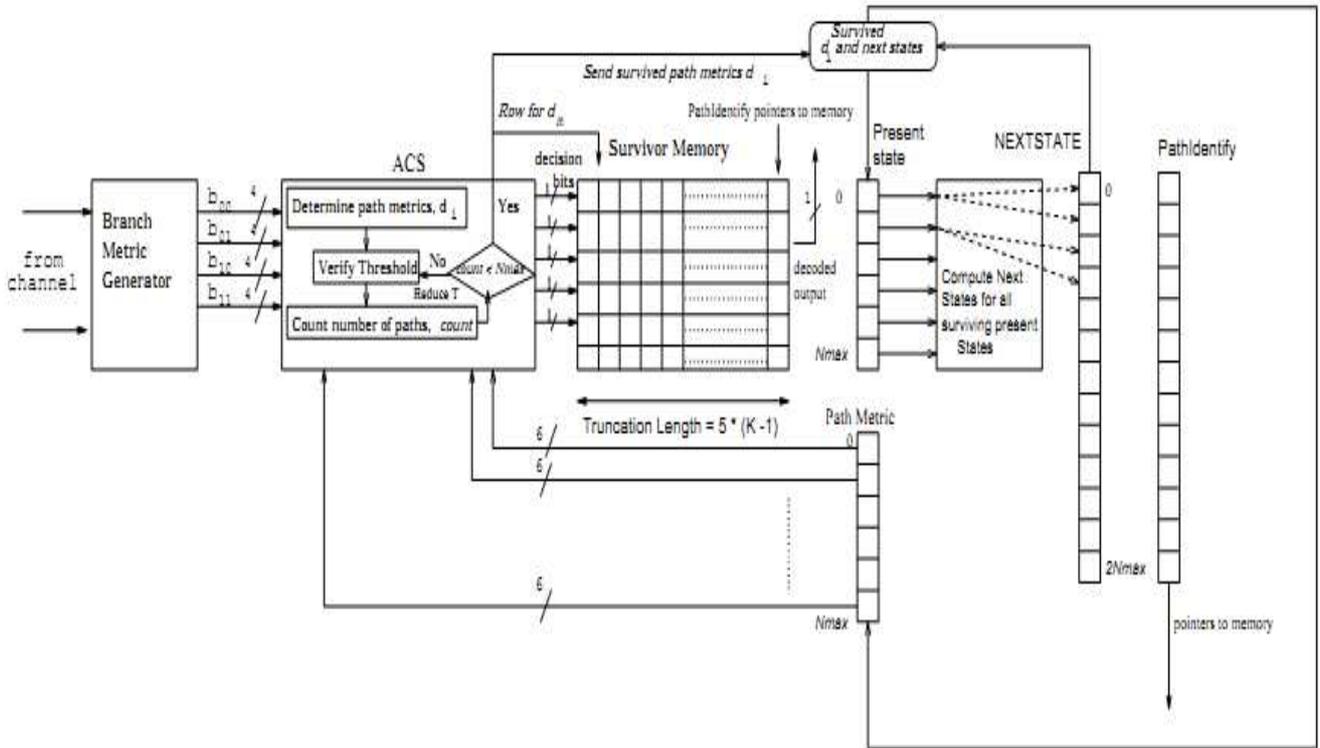


Fig.5 Adaptive Viterbi decoder architecture

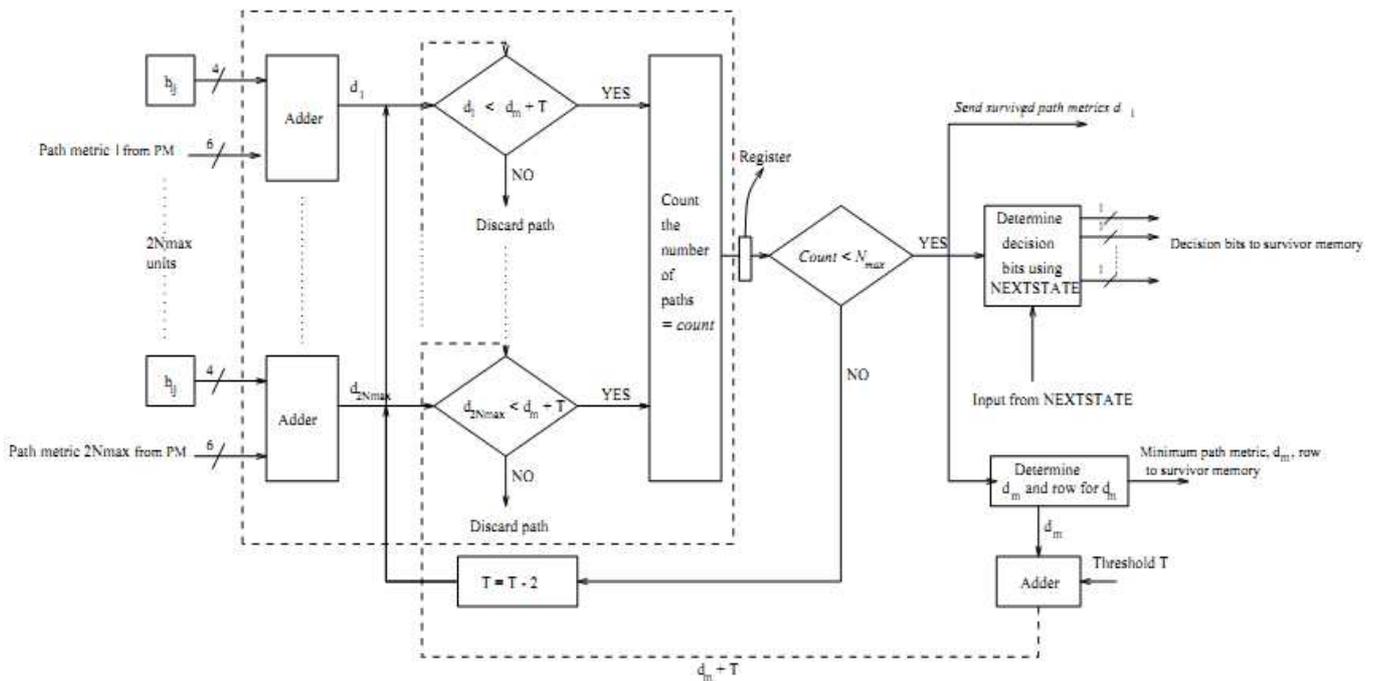


Figure 6: ACS unit of adaptive Viterbi decoder

The *Add-Compare-Select* (ACS) unit, shown in detail in Figure 6, evaluates the path metric of each path and determines if paths meet AVA conditions for path survival. At each trellis stage, the minimum-value surviving path metric among all path metrics for the preceding trellis stage, d_m , is computed. New path metrics are compared to the sum $d_m + T$ to identify path metrics with excessive cost. As shown at the left of Figure 6, the path metrics for all potential next state paths, d_i , are computed by the ACS unit. Comparators are then used to determine the life of each path based on the threshold, T . If the threshold condition is not satisfied by path metric $d_m + T$, the corresponding path is discarded.

Present and next state values for the trellis are stored in two column arrays, Present state and NEXTSTATE of dimensions N_{max} and $2N_{max}$ respectively, as shown in Figure 5. There can be at the most N_{max} survivor paths at any stage. Since each path is associated with a state, the number of present states is N_{max} . Each path can potentially create two child paths before pruning as there are two possible branches for each present state based on a received 0 or 1 symbol. Entries in the NEXTSTATE array need not be in the same row as their respective source present states. In order to correlate the next state paths and next states located in the NEXTSTATE array, an array of size $2N_{max}$, called PathIdentify, is used. For each next state element, this array also indicates the corresponding row in path storage (survivor) memory for the path.

Once the paths that meet the threshold condition are determined, the lowest-cost N_{max} paths are selected. To avoid the need for the sorting circuit described in [6] for the M-algorithm, we have developed a novel path pruning approach. Sorting circuitry is eliminated by making feedback adjustments to the parameter T . If the number of

paths that survive the threshold is less than N_{max} , no sorting is required. For stages when the number of paths surviving the threshold condition is greater than N_{max} , T is iteratively reduced by 2 for the current trellis stage until the number of paths surviving the threshold condition is equal to or less than N_{max} . In Section 7 it is shown that T and N_{max} can be determined through simulation so that T reduction is needed infrequently. Following path reduction, at most N_{max} remaining trellis states are stored in the *Present state* array in preparation for the receipt of the next symbol. The *register-exchange* based survivor memory [12] stores path sequence information and has a two dimensional size of $N_{max} * TL$, where TL is the truncation length. Each memory location stores an input bit, the *decision bit* from the ACS. Single-bit storage is performed for each surviving path at each trellis stage. Each row of the survivor memory is associated with a present state and has a *valid* bit to indicate the existence of a survivor path. Once survivor memory storage reaches the truncation length, the lowest-cost path sequence can be retrieved.

Architectural Model:

The logic area in terms of logic blocks of our adaptive Viterbi decoder can be characterized by an empirical model in terms of parameters N_{max} and K . This expression is of the form :

$$Area = AN_{max} + BKN_{max} + C$$

where A , B , C are constant coefficients. These coefficients were determined by evaluating a set of decoders with constraint lengths between 4 and 14. Through line-fitting, coefficients for A , B , and C were determined to be 90, 5.6, and 215, respectively.

The first term in the equation accounts for logic blocks which implement path metric comparators and the registers used to store path metrics. The second term accounts for path storage in memory implemented inside the FPGA. Since the truncation length (TL) is $5 * K$ [9], the term depends on both the maximum number of bits stored per trellis state, N_{max} , and K . Additionally, the width of present state and next state registers increases linearly with K and the number of registers increases linearly with N_{max} . The constant term in the expression accounts for logic which is fixed in size in relation to N_{max} and K . This includes thenbranch metric generator, which is dependent on the parameter ν defined in Section 2 and the logic needed to iteratively decrement T to avoid sorting.

Suitability of Dynamic Reconfiguration:

Dynamic reconfiguration of the entire adaptive Viterbi decoder hardware is considered as a means to enhance performance without compromising decode accuracy. The hardware resource requirements of an adaptive Viterbi decoder change in response to channel noise conditions for a fixed BER.

Based on noise levels at a specific time instant, minimally sufficient hardware resources can be dynamically allocated to meet the BER requirements of the application while achieving maximal performance [3]. A significant amount of channel noise demands a large constraint length, such as $K = 14$, to achieve a BER similar to that achieved by a constraint length $K = 4$ for a less noisy channel. It is shown in Section 5 that decoding speed is inversely related to resource requirement. This relationship is exploited to enhance performance through the dynamic allocation of AVA logic resources.

In implementing the AVA, two reconfiguration options, *fine-timescale* and *coarse-timescale* reconfiguration are considered. Coarse-timescale reconfiguration of the adaptive Viterbi decoder, based on parameters such as K , T and N_{max} , is performed in accordance to variations in channel noise conditions over seconds.

Reconfiguration at this time scale minimizes the performance impact of millisecond FPGA reconfiguration times. Coarse-timescale reconfiguration is motivated by changing channel noise characteristics from parameters such as weather, distance, or battery-power. These parameters result in a signal-to-noise ratio (SNR) that changes relatively slowly (seconds or longer).

When more accurate decoding is required, a lower clock-speed decoder (larger K) can be used at the cost of reduced decode rate. When less accurate decoding is required, a higher-performance decoder is swapped in. If dynamic reconfiguration was not allowed, the lower-performance decoder would always need to be resident. Coarse-timescale reconfiguration provides an optimized but variable bit rate and is targeted at data rather than voice applications.

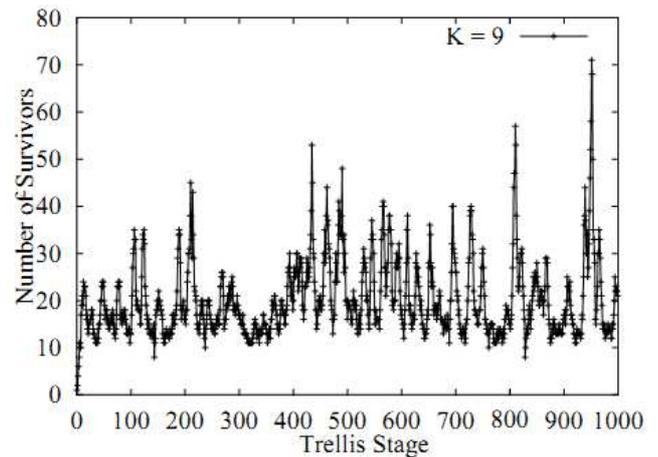


Fig.7: Number of surviving paths versus decoded symbol for the adaptive Viterbi algorithm for $K = 9$, $r = \frac{1}{2}$ convolutional code with N_{max} set to 2^{K-1} (its maximum) on an un-faded channel

In *fine-timescale* reconfiguration, configuration contents are changed once every one or two decoded bits (ms) in response to the number of survivor paths retained at specific instants of time. If additional survivor paths are required, the amount of required path storage increases, potentially limiting decoder performance. If fewer survivor paths are required, a faster-performance decoder can be used instead. After experimentation it was determined that fine-timescale reconfiguration is infeasible due to rapid variations (multiple changes per ms) in retained survivor paths over time. Current FPGA architectures [1] [15] require reconfiguration times measuring milliseconds. Figure 7 shows survivor path variation for a constraint length 9 decoder.

V. RESULTS AND DISCUSSIONS

Viterbi test Bench is created using Xilinx Webpack. Code is written in Verilog HDL. There are different Modules for the code and that are Viterbi test bench, Viterbi, Viterbi distance, reduce, path, path memory, compute metric, dff, ACS enable, back and compare select. To prove the correctness of our design, the verilog HDL description was simulated & tested using Model Sim Verilog Simulator. Afterwards Xilinx Webpack is used for design entry, synthesis, place & route and floor plan design. What is inputted to the test bench come from the Viterbi Encoder for $K=3$, rate = $\frac{1}{2}$ (111,101). For this Example the original data is: 010111001010001 + 00 (2 tail bits). The Encoder output is: 00 11 10 00 01 10 01 11 11 10 00 10 11 00 11 + 10 11. We may manually introduce error (s) into decoder input by altering the data (s).

Device selected is 2v40fg256-6.

Number of Slices: 86 out of 256 33%

Number of Slice Flip Flops: 62 out of 512 12%

Number of 4 input LUT's : 131 out of 512 25%

Number of bonded IOB's: 10 out of 88 11%

Number of GCLK's: 1 out of 16 6%

Figure 5.1 describes simulation waveforms of viterbi decoder. In this all values displayed are in nano seconds(ns). In this Timing numbers are only a synthesis estimate and trace report generated after place & route gives accurate timing information. Figure 5.2 gives Symbol of a Viterbi Decoder.From the figure it is clear that in this viterbi decoder there are six inputs & one output. The inputs are in0, in1, in2, in3, clk & reset & the output is decode out. Figure 5.3 gives RTL Schematic of Viterbi Decoder.

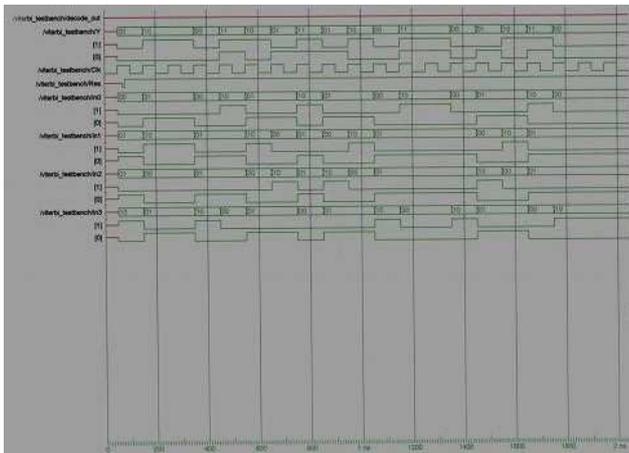


Fig.5.1 Simulation Waveform of Viterbi Decoder

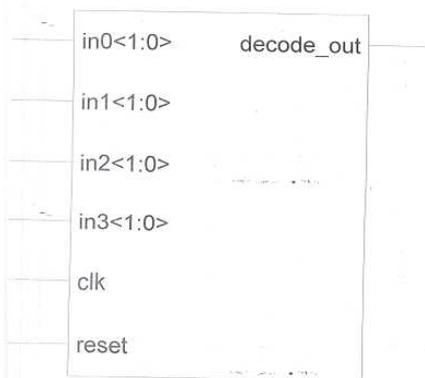


Fig.5.2 Symbol of Viterbi Decoder

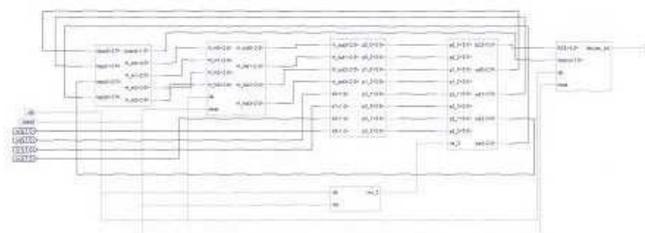


Fig.5.3 RTL schematic for Viterbi Decoder

VI. CONCLUSION AND FUTURE SCOPE

Viterbi Algorithm is widely used for the elimination of the potential noise in a data stream. Encoding is such that the Viterbi Decoder can remove potential noise in the incoming stream by decoding it. The characteristics of the

decoder are its effectiveness in noise elimination, speed of decoding and cost (hardware utilization). This thesis has presented the design and Implementation of the Viterbi Decoder. Its streamed input-output, regular architecture and parallel execution favor on an FPGA Implementation. FPGAs open a wide range of opportunities in the solution space that can result in high performance and economic solutions to a DSP problem because they do not map well to software programmable DSP architectures. The algorithm will have an ASIC solution but this may not be an option for reasons of schedule, economics of scale and flexibility. Applications such as data communications and image processing require more processing power but when the fastest DSP processor is not fast enough, the only alternatives are to add multiple DSP processors or to design custom hardware devices.

Multiple DSP processors are expensive, require many components and consume too much power. The performance gain that comes with each additional processor is small when compared to the increase in cost, board space, power consumption, and development time. Custom devices deliver the performance but sacrifice flexibility and require a large engineering investment with no chance to recover from mistakes. FPGAs are the new solution used by many engineers to implement computationally intensive algorithms.

FPGAs offer the best of all the worlds, the flexibility of a programmable solution, the performance of a custom solution, and lowering overall cost. So keeping in view of the advantages of using FPGAs as compared to other customized devices this work can further be increased to the designing & implementation of a generic Viterbi decoder. The generosity of the design facilitates not only the rapid prototyping of Viterbi decoders with different specifications but moreover, it explores the performance of different implementations in order to obtain the most suitable solution for a particular communication system. Some of the generic parameters are basic decoder specifications, metric size, trellis window length, number of surviving paths and pipeline depth. A new Viterbi decoder with new specifications can be realized by only re-synthesizing the code.

VII. REFERENCES

[1] Altera Corporation. *Apex II data sheet*, 2001.V <http://www.altera.com>.
 [2] Annapolis Microsystems, Inc. *WILD-ONE Reference Manual*, 1999.
 [3] W. Burluson, R. Tessier, D. Goeckel, S.Swaminathan, P. Jain, J. Euh, S. Venkatraman, and V. Thyagarajan. Dynamically Parameterized Algorithms and Architectures to Exploit Signal Variations for Improved Performance and Reduced Power. In *IEEE Conference on Acoustics, Speech, and Signal Processing*, May 2001.
 [4] F. Chan and D. Haccoun. Adaptive Viterbi Decoding of Convolutional Codes over Memoryless Channels.

IEEE Transactions on Communications, 45(11):1389–1400, Nov. 1997.

[5] M. Kivioja, J. Isoaho, and L. Vanska. Design and Implementation of a Viterbi Decoder with FPGAs. *Journal of VLSI Signal Processing*, 21(1):5–14, May 1999.

[6] C. F. Lin and J. B. Anderson. M-algorithm Decoding of Channel Convolutional Codes. In *Proceedings, Princeton Conference of Information Science and Systems*, pages 362–366, Princeton, NJ, Mar. 1986.

[7] A. Michelson and A. Levesque. *Error-control Techniques for Digital Communication*. John Wiley and Sons, New York, NY, 1985.

[8] B. Pandita and S. K. Roy. Design and Implementation of a Viterbi Decoder Using FPGAs. In *Proceedings, IEEE International Conference on VLSI Design*, pages 611–614, Jan. 1999.

[9] J. Proakis. *Digital Communications*. McGraw-Hill, New York, NY, 1995.

[10] H. Schmit and D. Thomas. Hidden Markov Modelling and Fuzzy Controllers in FPGAs. In *Proceedings, IEEE Workshop on FPGA-based Custom Computing Machines*, pages 214–221, Napa, Ca, Apr. 1995.

[11] S. J. Simmons. Breath-first Trellis Decoding with Adaptive Effort. *IEEE Transactions on Communications*, 38:3–12, Jan. 1990.

[12] S. Swaminathan. An FPGA-based Adaptive Viterbi Decoder. Master's thesis, University of Massachusetts, Amherst, Department of Electrical and Computer Engineering, 2001.

[13] R. Tessier and W. Bursleson. Reconfigurable Computing and Digital Signal Processing: A Survey. *Journal of VLSI Signal Processing*, 28(1):7–27, May 2001.

[14] Texas Instruments, Inc. *TMS320C6201 DSP Data Sheet*, 2001.

[15] Xilinx Corporation. *Virtex II data sheet*, 2001. <http://www.xilinx.com>.

[16] D. Yeh, G. Feygin, and P. Chow. RACER: A Reconfigurable Constraint-Length 14 Viterbi Decoder. In *Proceedings, IEEE Workshop on FPGA-based Custom Computing Machines*, Napa, Ca, Apr. 1996.

About Authors:



Shoban Mude is an Pursuing M.Tech Post-Graduate of VLSI-SD from Guru Nanak Institute of Technology, Department of ECE, JNTUH. he obtained his B.Tech from S R College of Engineering and Technology, Warangal. he Has 1.6 Years of Teaching Experience. His interesting Fields are VLSI, and Communications.



S Nagakishore Bhavanam is an M.Tech Post-Graduate of VLSI-SD from Aurora's Technological & Research Institute (ATRI), Department of ECE, JNTUH. He obtained his B.Tech from S.V.V.S.N Engineering college, Ongole. He has 2.6 Years of teaching experience. He has 4 Research Papers, Published in IEEE Xplore, He has 4 International journal Publications and has 6 Papers, Published in International & National Conferences. His interesting Fields are Low Power VLSI, Digital System Design, Sensor Networks, and Communications.

Evaluation of Shear Strength of SFRC Moderate Deep Beams Using Strut-and-Tie Models

Vinu R. Patel¹, Dr. I. I. Pandya²

¹Assistant Professor, Department of Applied Mechanics, Faculty of Technology & Engineering, M.S. University, Baroda, 390020, Gujarat, India.

²Associate Professor, Department of Applied Mechanics, Faculty of Technology & Engineering, M.S. University, Baroda, 390020, Gujarat, India.

ABSTRACT

Strut-and-tie is a system of forces' distribution in the form of "load-path" connectivity from the applied load point to the support point. Strut-and-Tie Method (STM) has been developed based on simple truss model. STM models represent the load carrying mechanism of a structural member by approximating the flow of internal forces by means of struts representing the flow of compressive stresses and ties representing the flow of tensile stresses. IS 456: 2000, along with other various codes of different countries, classifies the beam into three categories; namely normal beam, moderate deep beam, and deep beam, according to their span to depth ratios. The aim was to provide a systematic and comprehensive study on the shear strength of SFRC (Steel Fiber Reinforced Concrete) moderate deep beams without web reinforcement and to compare experimental result of shear strength with theoretical result by STM of SFRC. Experimental results of ultimate shear strength are compared with the theoretical results calculated from formula of STM given by various source such as ACI 318-08, Nielsen (1984), and Schlaich et al. (1987). We found that the experimental value and theoretical value by STM are within 15% variation range for all types of beam.

Keywords: Strut-and-tie, Steel Fiber Reinforced Concrete, Moderate deep beam, Ultimate shear strength, stress-strain.

1. Introduction

Strut-and-Tie Method (STM) has been developed based on simple truss model. The simple truss model is only rational for the design of cracked reinforced concrete beams. The design based on the simple truss model is limited to certain parts of structure. However, STM as an extension to simple truss model is applicable to analyze and design the whole of a reinforced concrete member experiences three effects; shear, flexural, and axial effect. The original strut-and-tie model has been developed as a Lower Bound solution of plastic theory where equilibrium of a system is considered together with the yield criterion.

Since 2002, strut-and-tie method has been included as an alternative design method in North America [ACI 318-2008]^[2]. This report describes the development of strut-and-tie theory in analysis of fibrous reinforced concrete beam. A shear behavior of moderate deep beams is included as an example, and compared theoretical values with experimental values. The scope of this study covers the development of the strut-and-tie method as an analytical procedure based on Appendix A of ACI 318-2008^[2] and other sources.

The concept of incorporating strong thin fiber to strengthen brittle matrices is not new. The concept is more than 4500 years old. Potentially, the addition of fibers causes substantial

changes in properties of both fresh and hardened concrete. Due to low effectiveness, poor alkaline resistance high cost, use of other fibers such as nylon, rayon, carbon etc. has been almost ruled out after initial investigation. The use of strong and stiff fibers in concrete improves the post cracking performance of concrete considering reserved strength. After micro cracking, fibers spanning the cracks, control crack propagation and control the rate of widening of cracks under tensile loading. This role of fiber imparts ductility of concrete and delays its failure. The process of fiber pull out absorbs lot of energy and hence the toughness of concrete and its impact resistance are considerably increased. Remarkable improvements in elastic modulus, tensile strength, crack resistance, crack control, durability, fatigue resistance, impact resistance, abrasion resistance etc., resulted in FRC material, which arrived as a boon to overcome the drawbacks of steel-reinforced concrete.

2. Research Signification

The scope of this study covers the development of the strut-and-tie method as a design procedure and to compare ultimate shear strength of fiber reinforced concrete moderate deep beams without web reinforcement (Stirrups) using strut-and-tie method with experimental results.

Table 1: Formulas given by Different Sources

| Sources | Strut compressive capacity |
|------------------------|--|
| ACI 318-08 | <p>Without Longitudinal Reinforcement</p> $0.85\beta_s f'_c A_{cs}$ Prismatic: $\beta_s = 1.0$ Bottle-Shaped w/reinf. satisfying crack control: $\beta_s = 0.75$ Bottle-Shaped not satisfying crack control: $\beta_s = 0.60\lambda$ $\lambda = 1.0$ for normal weight concrete $\lambda = 0.85$ for sand-lightweight concrete $\lambda = 0.75$ for all lightweight concrete Strut in tension members: $\beta_s = 0.40$ All other cases: $\beta_s = 0.60$ |
| Schlaich et al. (1987) | <p>With Longitudinal Reinforcement</p> $f_{cu} A_c + f'_s A'_s$ $0.85f'_c$ “for an undisturbed and uniaxial state of compressive stress” (prismatic) $0.68f'_c$ “if tensile strains in the cross direction or transverse tensile reinforcement may cause cracking parallel to the strut with normal crack width” $0.51f'_c$ “as above for skew cracking or skew reinforcement” $0.34f'_c$ “for skew cracks with extraordinary crack width. Such cracks must be expected, if modeling of the struts departs significantly from the theory of elasticity’s flow of internal forces” |
| Nielsen (1984) | $(0.8 - \frac{f'_c}{200}) f'_c A_{cs}$ |

Table 2: Notations for Table 1

| Notations |
|--|
| A'_s = area of compression steel (in ²) |
| A_c = area of concrete in the strut (in ²) |
| A_{cs} = area of concrete in the strut (in ²) |
| A_{si} = total area of surface reinforcement at spacing s_i (in ²) |

f_c = concrete compressive strength (ksi)
 f_{cu} = effective concrete compressive strength (ksi)
 α_i = the angle between the reinforcement and the axis of the strut (DEG.)

3. Experimental Programme

Testing was carried out on 12 SFRC using Flat Corrugated Type (FCT) steel fibers moderate deep beams. These beams were tested in simply supported conditions under two equal point loading each at a distance of 1/3 of effective span from support.

3.1 Test specimen

Twelve Steel Fiber Reinforced Concrete moderate deep beams, simply supported on effective span of 1200 mm were tested under one point loading. Length of the beams and width of the web were kept constant (1300 mm and 150 mm respectively). The beams were divided into four series having depths of 300 mm, 400 mm, 500 mm and 600 mm. Each series comprised of three beams. i.e. beam notation "D60" denotes the beam having overall depth D of 60 cm.

3.2 Test materials

The cement used was ordinary Portland cement of grade 53. ordinary river sand having fineness modulus of 2.8 and maximum size of 4.75 mm, and crushed basalt gravel having a maximum size of 20 mm were used as a fine and coarse aggregate respectively. The concrete mix proportion was 1:1.5:3 (cement: fine aggregate: coarse aggregate) by weight with flat corrugated type steel fiber volume fraction of 1 % by volume of concrete and water cement ratio of 0.45 kept constant for all beams. Longitudinal tension reinforcement consists of High yield strength deformed bars (415 N/mm²) used, Vertical Shear Reinforcement (stirrups) are not provided. Six cubes (150mm) and eight cylinders (four cylinders for compressive strength and four cylinders for splitting strength, 150mm diameter and 300mm height) were cast as control specimens from each mix. All specimens were cured at least for 28 days.

3.3 Testing procedure

All the beams were tested under two point concentrated loadings positioned at one third span. All the beams were simply supported with an effective span of 1200 mm. Beams were centered on platform and leveled horizontally using level tube and vertically by adjusting the bearing plates. Load was applied gradually.

Crack propagations were traced by pencil and their tips were marked corresponding to the load readings.

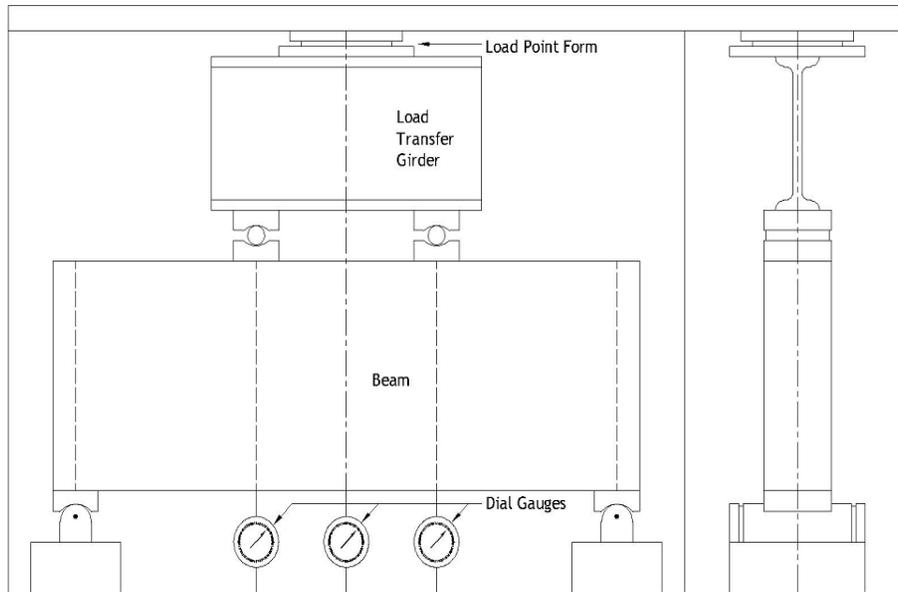


Fig.1: Test setup

4. Discussion of Test Result

Experimental results are compared with theoretical results. Theoretical results are calculated from formula of STM which are given by various sources such as ACI 318-08^[2], Nielsen (1984)^[3], and Schlaich et al. (1987)^[5,6]. The results of 12 SFRC (FCT) compared with theoretical results.

Table 3 Comparison of $V_{u(exp)}$ and $V_{u(th)}$ for SFRC (FCT)

| | Experimental Result | ACI 318-08 | Nielsen (1984) | Schlaich et al. (1987) |
|------------|---------------------|------------|----------------|------------------------|
| D30 | 7.55 | 7.073 | 7.422 | 7.544 |
| | 8.00 | 7.258 | 7.576 | 7.742 |
| | 7.05 | 7.021 | 7.379 | 7.489 |
| | 13.00 | 11.705 | 12.285 | 12.485 |
| D40 | 14.00 | 12.863 | 13.226 | 13.720 |
| | 13.45 | 11.807 | 12.370 | 12.594 |
| | 17.00 | 15.456 | 16.010 | 16.486 |
| D50 | 17.25 | 15.704 | 16.209 | 16.751 |
| | 17.85 | 15.493 | 16.040 | 16.526 |
| | 23.60 | 20.512 | 20.930 | 21.880 |
| D60 | 23.70 | 20.903 | 21.231 | 22.296 |
| | 22.65 | 19.795 | 20.369 | 21.115 |

Table 4 Ratio of ($V_{u(exp)}/V_{u(th)}$) for SFRC (FCT)

| | a/h | l/h | ACI 318-08 | Nielsen (1984) | Schlaich et al. (1987) |
|------------|------|-----|------------|----------------|------------------------|
| D30 | 1.33 | 4 | 1.067 | 1.017 | 1.001 |
| | 1.33 | 4 | 1.102 | 1.056 | 1.033 |
| | 1.33 | 4 | 1.004 | 0.955 | 0.941 |
| D40 | 1 | 3 | 1.111 | 1.058 | 1.041 |
| | 1 | 3 | 1.088 | 1.059 | 1.020 |
| | 1 | 3 | 1.139 | 1.087 | 1.068 |
| D50 | 0.80 | 2.4 | 1.100 | 1.062 | 1.031 |
| | 0.80 | 2.4 | 1.098 | 1.064 | 1.030 |
| | 0.80 | 2.4 | 1.152 | 1.113 | 1.080 |
| D60 | 0.66 | 2 | 1.151 | 1.128 | 1.079 |
| | 0.66 | 2 | 1.134 | 1.116 | 1.063 |
| | 0.66 | 2 | 1.144 | 1.112 | 1.073 |

*Unit of V_u is **Tons (UK)**.

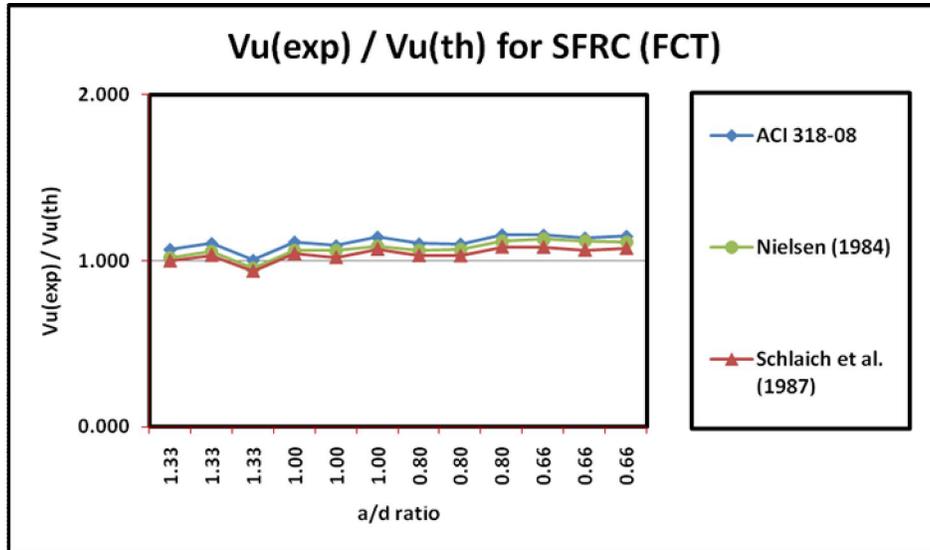


Fig. 2: Graphical presentation of $V_{u(\text{exp})}/V_{u(\text{th})}$ ratio For SFRC (FCT) beams

5. Conclusion

- I. Experimental results of ultimate shear strength are compared with the theoretical results calculated from formula of STM given by various source such as ACI 318-08^[2], Nielsen (1984)^[3], and Schlaich et al. (1987)^[5,6]. The Table 3 & Table 4 indicates indirect verification of experimental results with theoretical results. The theoretical results by STM are within $\pm 15\%$ variation for all types of beam.
- II. The average ratio of ($V_{\text{exp}}/V_{\text{th}}$) for beams of D30 series is 1.058 for ACI 318-08^[2], 1.009 for Nielsen (1984)^[3], and 0.992 for Schlaich et al. (1987)^[5,6].
- III. The average ratio of ($V_{\text{exp}}/V_{\text{th}}$) for beams of D40 series is 1.113 for ACI 318-08^[2], 1.068 for Nielsen (1984)^[3], and 1.043 for Schlaich et al. (1987)^[5,6].
- IV. The average ratio of ($V_{\text{exp}}/V_{\text{th}}$) for beams of D50 series is 1.117 for ACI 318-08^[2], 1.080 for Nielsen (1984)^[3], and 1.047 for Schlaich et al. (1987)^[5,6].
- V. The beams of series of D60 average of ratio of ($V_{\text{exp}}/V_{\text{th}}$) is 1.143 for ACI 318-08^[2], 1.119 for Nielsen (1984)^[3], and 1.072 for Schlaich et al. (1987)^[5,6].
- VI. This shows that Schlaich et al. (1987)^[5,6] and Nielsen (1984)^[3] predict conservative shear strength, but nearly accurate results for all beams.
- VII. The ACI 318-08^[2] gives very conservative results for all beams.

6. Nomenclature

ACI : American Concrete Institute
STM : Strut-and-Tie Models or Strut-and-Tie Method
SFRC : Steel Fiber Reinforce Concrete
FRC : Fiber Reinforce Concrete
 V_u : Ultimate Shear Strength
FCT : Flat Corrugated Type

7. References

1. IS 456: 2000, Plain and Reinforced Concrete — Code of Practice, Fourth Revision, Bureau of Indian Standards (BIS 2000), Fifth Reprint August 2002).
2. ACI 318-2008, Building Code Requirements for Structural Concrete and Commentary, Appendix A, Strut-and-Tie Models, American Concrete Institute, Farmington Hills.
3. Nielsen M. P., Limit Analysis and Concrete Plasticity, Prentice-Hall, Englewood Cliffs, New Jersey, USA.
4. Nielsen, M. P. and Braestrup, M. W., Shear Strength of Prestressed Concrete Beams without Web Reinforcement, "Magazine of Concrete Research", 30, 104, 1978, pp 119-128.
5. Schlaich J. and Schafer K., Design and Detailing of Structural Concrete Using Strut-and-Tie Models, "The Structural Engineer". V. 69, No. 6, May-June, 1991, pp 113-125.
6. Schlaich J., Schaefer K., and Jennewein M., Toward a Consistent Design of Structural Concrete, "PCI Journal", Vol. 32, No. 3, May-Jun 1987, pp 74-150.

Measurement of Plasma Resistivity in the Edge of IR-T1 Tokamak

A. Razmara¹, P. Khorshid¹, M. Ghoranneviss² and H. Arabshahi³

¹Department of Physics, Islamic Azad University, Mashhad Branch, Mashhad, Iran

²Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran

³Department of Physics, Payame Nour University of Fariman, Fariman, Iran

Abstract— The Plasma resistivity has been measured in the edge plasma of IR-T1 tokamak. To achieve the edge plasma properties (T_e and n_e) using I-V characteristics of movable single langmuir probe. Density and temperature of electron, respectively n_e and T_e , are measured simultaneously for calculating the radial profile of parallel and vertical resistivity. As temperature of plasma raised, resistivity drops rapidly so Reduction of resistivity with increasing temperature could leads to prolonging of the duration of the plasma discharge.

Keywords-: tokamaks; resistive MHD; Electron collisions

I. INTRODUCTION

Any realistic plasma will have a density gradient, and the plasma will tend to diffuse toward regions of low density. We assume that the plasma is weakly ionized, so that charge particles collide primarily with neutral atoms rather than with one another. As the plasma spreads out as a result of pressure gradient and electric field forces, the individual particles undergo a random walk, colliding frequently with the neutral atoms. When plasma consists of just electrons and ions, all collisions are coulomb collisions among charge particle (Trintchouk et al., 2003)(Chen, 1984). If a uniform steady electric field is imposed on plasma this electric field will accelerate the ions and electrons in opposite directions. The accelerated particles will collide with other particles and this fractional drag will oppose the acceleration. Resistivity is determined by the collisional drag on electrons moving against the background of ions. Suppose that an electric field E exists in a plasma and that the current that it drives is all carried by the electrons, which are much more mobile than the ions. Then in steady state the electron equation of motion changes, so that

$E = \eta J$ is ohm's low, and η is the resistivity. The transverse or cross field resistivity was calculated by Spitzer as the rate of momentum transfer from electrons to ions through collisions in a resistive magnetohydrodynamics (Trintchouk et al., 2003)(de Blank, 2006) ,

where the electron temperature T_e is in electron volts, Z is the ion atomic mass and $\ln \Lambda$ is the coulomb logarithm. Also parallel resistivity defined by (Mahmoodi Darian et al., 2006)(Goldstone, 1995) ,

$$\eta_{\perp} \equiv 1.03 \times 10^{-4} T_e^{-3/2} Z \ln \Lambda \text{ (ohm.m)} \quad (1)$$

$$\eta_{Spitzer} = \eta_{\parallel} = 5.24 \times 10^{-5} \frac{Z \ln \Lambda}{T_e^{3/2}} \text{ (ohm.m)} \quad (2)$$

Impurities in plasma are one of important factors of instabilities. For example impurities prevent plasma from heating. We have defined Z_{eff} as the ratio of the measured plasma resistivity η_p to the theoretical resistivity η_{\parallel} . For ohmic input power, plasma resistivity is,

$$\eta_p = \frac{r^2 V_l}{2RI_p} \quad (3)$$

where V_l is loop voltage, I_p is the plasma current and R is the major radius of tokamak chamber. By definition of Z_{eff} the value of η_p could be calculated. The time evolution of Z_{eff} from a typical discharge of IR-T1 Tokamak has been shown in Figure 1.

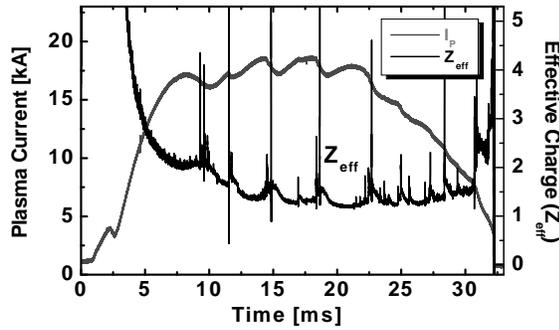


Figure 1. Time evolution of plasma current and effective charge.

II. EXPERIMENTAL SET-UP AND DIAGNOSTICS

IR-T1 tokamak is an ohmically heated air core tokamak with a major radius of $R=0.45$ m and a minor radius of $a=0.125$ m defined by two poloidal stainless-steel limiters. The vacuum chamber has a circular cross-section with two toroidal breaks and a minor radius of $b=0.15$ m. Toroidal magnetic field is equal to $B_t \sim 0.6-0.8$ T, plasma current is $I_p \sim 25-30$ kA, averaged electron density in hydrogen is $0.5-1.5 \times 10^{19} \text{ m}^{-3}$, plasma discharge duration is $t_d \sim 30$ ms and electron temperature is $T_e(0) \sim 150-180$ eV. A single Langmuir probe was used to measure spatial and temporal evolutions of electron temperature and density. The probe is connected to the power supply where its potential is varied continuously over a range from negative to positive potential with respect to the plasma potential to obtain the ion and electron current. The current, I is determined as a function of the applied probe voltage V_{app} . This relation $I = f(V_{app})$ is called the probe characteristic or the $I-V$ characteristic. The general appearance of the $I-V$ characteristic for a Langmuir probe is shown in Figure 2.

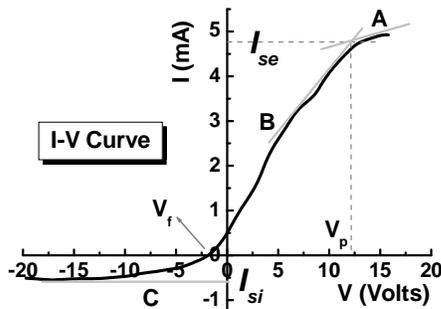


Figure 2. The $I-V$ characteristic Langmuir probe during a typical discharge of IR-T1 tokamak.

In the Figure 2. the region A is known as electron saturation and I here is equal to the electron saturation current, I_{se} ,

$$I_{es} = \frac{1}{4} en_e V_{e,th} A \quad (4)$$

where n_e is the electron density, $V_{e,th} = \sqrt{8kT_e / \pi m_e}$ is the electron thermal speed. If the potential of probe becomes less than V_p , $V < V_p$, the probe is negative with respect to the surrounding plasma and this causes to reflect part of the impacting electrons (region B). Eventually in this potential, I reduce to small fraction of saturation current. The total current is zero when $I_i \approx I_e$. This phenomena happen when the probe potential is equal to floating potential V_f . Decreasing the potential further (entering region C) probe with constant rate, enable to collected ions. This is ion saturation current, $I = I_{si}$,

$$I_{is} = \frac{1}{4} en_i V_{i,th} A \quad (5)$$

where n_i is the ion density $V_{i,th} = \sqrt{8kT_i / \pi m_i}$ is the ion thermal speed and A is the probe collecting area. When $T_e \gg T_i$, the ion saturation current is not determined by the ion thermal speed, and determined by the Bohm ion current (Merlino, 2007)(Popescun and Ohtsu, 2007),

$$I_{is} = I_{Bohm} = 0.6 en_i \sqrt{\frac{kT_e}{m_i}} A \quad (6)$$

From above equation, it could be seen that because of $n_e = n_i$ and $m_e \ll m_i$, the electron saturation current will be much greater than the ion saturation current (Merlino, 2007). Radial profile of electron density and electron temperature has been calculated by Langmuir probe, behind and in front of the fixed poloidal limiter from $r = 10.5-13.5$ cm. The results have been shown in Figure 3.

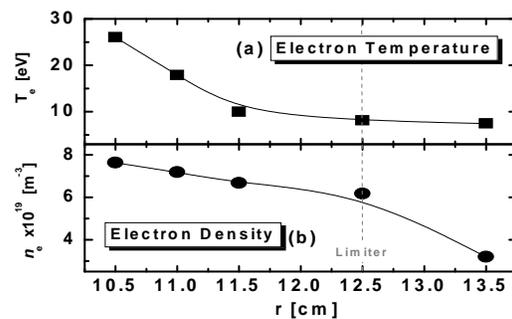


Figure 3. (a) The radial profile of electron temperature and (b) the radial profile of electron density.

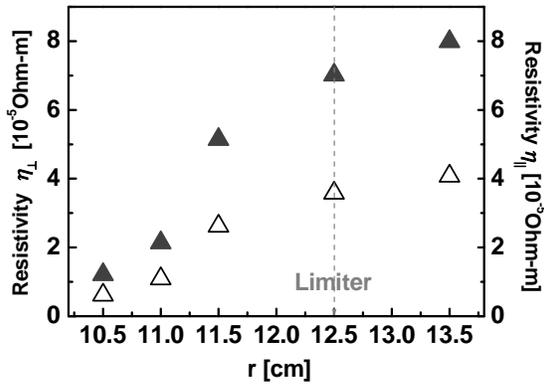


Figure 4. The radial profile of parallel and vertical plasma resistivity.

The radial profile of electron temperature and electron density have been shown in Figure 3. it can deduced that the electron temperature and electron density are increasing from edge to core of plasma, so according to equation 2 the plasma resistivity will decrease by increasing electron temperature. In other hand, when electron temperature increases, the collision frequency is decrease. Collisions between electrons and ions in plasma will prevent the accelerations of electrons in response to an electric field. Without such collisions, electrons would be accelerated indefinitely by an applied electric field, so that an infinitesimal voltage would be sufficient to drive a large current through plasma. Collisions between electrons and ions cause to limit the current that can be driven by an electric field. Also resistivity may be expressed in terms of the electron-ion collision frequency, (Goldstone, 1995)

$$v_{ei} = \left(\frac{ne^2}{m}\right)\eta \quad (7)$$

So we can deduce from above sentences that when electron temperature increases, collision frequency decrease therefore plasma resistivity decreases. The Figure 4. shows the calculated radial profile of parallel and vertical plasma resistivity it decreased from edge to core. We found that, transverse resistivity for electrons are higher than the parallel resistivity by a factor of 1.96, the results have been compared with other tokamak plasmas and it is similar to TCABR and CASTOR tokamak. In the last, the results have shown that electron density increased from edge to core of plasma column, as the parallel electrical resistivity of the plasma decreases with increasing temperature, this could leads to prolonging of the duration of the plasma discharge.

III. CONCLUSION

In this paper we obtained the plasma resistivity profile using parameters of plasma such as n_e and T_e calculated by I - V characteristics of moveable single Langmuir probe in the edge of IR - TI tokamak. The results showed that electron temperature decrease from core to edge, as for electron density is. Plasma resistivity is proportional to $T_e^{-3/2}$, so resistivity increase from center to edge of plasma. As temperature of plasma raised, resistivity drops rapidly so plasma at very high temperatures is collision-less i.e. their resistivity is negligible. Reduction of resistivity with increasing temperature could leads to prolonging of the duration of the plasma discharge. Effective charge may be important to measurement of resistivity in high plasma impurity, so that it can prevent the plasma for ohmically heating.

IV. ACKNOWLEDGMENTS

The authors would like gratefully to acknowledge IR-T1 team for their technical assistance and author A.R. thanks Dr. Z. Emami for her helpful discussions.

REFERENCES

- [1] F. F. Chen, (1984). Introduction to Plasma Physics and Controlled Fusion, Plenum Press, Los Angeles.
- [2] H. J de Blank, (2006), Transactions of Fusion Science and Technology, **49**, 118.
- [3] R. J. Goldstone, (1995). Introduction to Plasma Physics, IOP Publishing co., Bristol and Philadelphia.
- [4] M. Mahmoodi Darian, A. Hojabri and M. Salem, (2006), *Journal of Fusion Energy* **25**, 241.
- [5] R. L. Merlino, (2007). J. Phys, **75**, 1078.
- [6] S. Popescun, Y. Ohtsu, (2007). Appl. Phys, **102**. 093302.
- [7] F. Trintchouk, M. Yamada, H. Ji, R. Kulsrud and T. Carter, (2003). Physics of Plasmas, **10**, 319.

EFFECT OF COUPLE-STRESS ON THE REFLECTION AND TRANSMISSION OF PLANE WAVES AT AN INTERFACE

¹Mahabir Barak, ²Anil Kumar

¹Department of Mathematics, Pt. NRS.Govt. College,Rohtak, Haryana (India)

²Department of Computer Science, Pt. NRS.Govt.College,Rohtak, Haryana
(India)

ABSTRACT

The present paper is concerned with the reflection and transmission of plane waves at an interface between two dissimilar couple-stress elastic half-spaces in perfect contact. Amplitude ratios of various reflected and refracted waves have been calculated and computed numerically for a specific model. The variations of these amplitude ratios with the angle of incidence have been shown graphically. The results of Graff and Pao [3] have been obtained as a special case and are shown graphically. The amplitude ratios at elastic-elastic interface have also been deduced.

Key Words: Couple–stress, elastic media, Amplitude ratios, Angle of incident, Welded contact, Stress free boundary.

INTRODUCTION

In the development of classical theory of elasticity the potential energy density is assumed to be a function of strain components, which are formed by a linear combination of the first order space derivatives of the displacements. The effects of surface moment per unit area have been neglected in the theory of elasticity [4]; and only the surface force per-unit

area, known as traction, has been taken into account. The components of traction are the force-stresses, extensions of the classical theory has been made to include the effects of higher-order derivatives of displacement in the potential energy. The work done by surface couples, surface forces and higher order stress like quantities have been included in the balance of energy in the new developments.

The development in the field of couple-stresses was mainly done by Mindlin and Tiersten [5], they developed a linear theory with a complete set of equations of motion, constitutive equations and boundary conditions for unique solutions. Aggarwal and Alverson [1] studied the effect of couple-stresses on diffraction of plane elastic waves by cylindrical discontinuities.

Sengupta and Ghosh [6] studied the effect of couple-stresses on the surface waves in elastic media, they deduced the equations of surface waves in elastic media under the influence of couple stresses and observed that the effect of couple-stresses increases the velocity of Rayleigh and Love wave propagation. Sengupta and Benerji [7] investigated the effects of couple-stresses on propagation of waves in elastic layer immersed in an infinite liquid. Graff and Pao [3] discussed the effects of couple-stresses on the propagation and reflection of plane waves in an elastic half space and obtained theoretically the amplitude ratios of various reflected waves. Recently Singh and Kumar [8] studied the reflection and refraction of plane waves at an interface between micropolar elastic solid and viscoelastic solid. Kumar and Singh [9] investigated the reflection of plane waves at a planer viscoelastic micropolar interface. Singh and Kumar [10] discussed the wave propagation in a micropolar elastic solid with stretch. In the present investigation, the problem of reflection and transmission of plane wave at an interface between two dissimilar couple-stress elastic solids in

welded contact has been discussed. The results of Graff and Pao [3] have been deduced as a special case and are presented graphically.

BASIC EQUATION AND THEIR SOLUTIONS

The equation of motions without body forces and body couples and the constitutive relations in couple stress theory of elasticity are given by Mindlin, and Tiersten (1962) as

$$\mu \nabla^2 \bar{u} + (\lambda + \mu) \text{grad div } \bar{u} + \eta \nabla^2 (\text{grad div } \bar{u}) - \eta \nabla^2 (\nabla^2 \bar{u}) = \rho \frac{\partial^2 \bar{u}}{\partial t^2}, \quad (1)$$

$$\mathbf{T}_{ji} = \mathbf{T}_{ji}^s + \frac{1}{2} \epsilon_{j\bar{i}} (\mu_{p',p}^D + \mu_{s,l}), \quad (2)$$

with

$$\mathbf{T}_{ji}^s = \lambda \epsilon_{kk} \delta_{ij} + 2 \mu \epsilon_{ij} \quad (3)$$

$$\mu_{ij} = \mu_{ij}^D + \mu \delta_{ij} \quad (4)$$

where

$$\mu_{ij}^D = 4\eta \Psi_{ij} + 4\eta' \Psi_{ji} \quad (5)$$

$$\epsilon_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i})$$

$$\Psi_{ij} = \omega_{j,i} = \frac{1}{2} \epsilon_{jkl} u_{l,ki} \quad (6)$$

and

$$\mu_{ij}^D = \mu_{ij} - \mu \delta_{ij}.$$

The list of symbols is given at the end of the paper.

By Helmholtz representation of displacement vector, we can write

$$\bar{u} = \nabla \phi + \text{curl } \bar{\psi} \quad \nabla \cdot \bar{\psi} = 0 \quad (7)$$

Making use of equation (4) in equation (1), we get

$$\nabla^2 \phi = \frac{1}{C_1^2} \frac{\partial^2 \phi}{\partial t^2} \quad (8)$$

and

$$\nabla^2 (1 - l^2 \nabla^2) \psi = \frac{1}{C_2^2} \frac{\partial^2 \psi}{\partial t^2} \quad (9)$$

where

$$C_1^2 = \frac{\lambda + 2\mu}{\rho}, \quad C_2^2 = \frac{\mu}{\rho}, \quad l^2 = \frac{\eta}{\mu}$$

where ψ is the zth component of $\vec{\psi}$ and $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$

We are considering two dimensional problems in xy-plane, therefore the displacement and rotation vectors are taken as

$$\begin{aligned} \vec{u} &= (u_x, u_y, 0) \\ \vec{w} &= (0, 0, w_z) \end{aligned} \quad (10)$$

Assuming time harmonic variation as $\exp(-i\omega t)$ in equation(5)and (6),we obtain

$$(\nabla^2 + \delta_1^2)\phi = 0 \quad (11)$$

$$(-l^2 \nabla^4 + \nabla^2 + \delta_2^2) \psi = 0, \quad (12)$$

where $\delta_1^2 = \frac{\omega^2}{C_1^2}$ $\delta_2^2 = \frac{\omega^2}{C_2^2}$

To solve equation(11) and (12),and we assume the solution of equation (12) as,

$$\psi = \psi_1 + \psi_2$$

where ψ_1 and ψ_2 satisfy the following equations

$$(\nabla^2 + \delta_2^2)\psi_1 = 0, \quad (\nabla^2 + \delta_3^2)\psi_2 = 0$$

$$\phi = f(y) e^{i(\xi x - \omega t)} \quad (10)$$

$$\psi = g(y) e^{i(\zeta x - \omega t)} \quad (11)$$

where ξ and ζ are the wave numbers. Substitution of (10) and (11) into (8) and (9) ,Yield second order and fourth order differential equation and the fourth order differential equation can be reduced to, two second order differential by letting

$$\psi(y) = \psi_1(y) + \psi_2(y) \quad (12)$$

Then the equations are

$$\frac{d^2f}{dy^2} + \alpha^2 f = 0 \quad (13)$$

$$\frac{d^2\psi_1}{dy^2} + \beta^2 \psi_1 = 0 \quad (14)$$

$$\frac{d^2\psi_2}{dy^2} - \gamma^2 \psi_2 = 0 \quad (15)$$

where

$$\alpha^2 = \frac{\omega^2}{c_1^2} - \xi^2 \quad (16)$$

$$\beta^2 = \beta_1^2 - \zeta^2 \quad (17)$$

$$\gamma^2 = \beta_2^2 + \zeta^2 \quad (18)$$

and

$$\beta_1^2 = \left\{ 1 + \sqrt{1 + \frac{4\ell^2\omega^2}{C_2^2}} \right\} / 2\ell^2,$$

$$\beta_2^2 = \left\{ 1 - \sqrt{1 + \frac{4\ell^2\omega^2}{C_2^2}} \right\} / 2\ell^2 \quad (19)$$

From equations (16)-(18), it is seen that γ^2 is always greater than zero, where as α^2 and β^2 may have real value greater than equal to or less than zero. The solutions of (13) and (14) will of course differ accordingly.

Therefore, in an unbounded couple stress medium three types of waves propagate. In addition to P and SV-wave, there is SS-wave which decays rapidly from the surface for any reasonable value of the couple-stress material constant.

The stresses in terms of potentials are

$$T_{xy} = \mu \left[\frac{2\partial^2\phi}{\partial x \partial y} + \frac{\partial^2\psi}{\partial y^2} - \frac{\partial^2\psi}{\partial x^2} - l^2 \nabla^4 \psi \right], \quad (20)$$

$$T_{yy} = \mu \left[K_1^2 \nabla^2 \phi - \frac{2\partial^2\phi}{\partial x^2} - \frac{2\partial^2\psi}{\partial x \partial y} \right], \quad (21)$$

$$\mu_{yz} = -2\eta \frac{\partial}{\partial y} (\nabla^2 \psi), \quad (22)$$

where

$$K_1^2 = C_1^2 / C_2^2$$

III. FORMULATION OF THE PROBLEM

We consider two homogeneous couple-stress elastic half spaces M and M' which are in welded contact at the interface $y = 0$. We take y-axis vertically downward and x-axis horizontally. We consider a plane harmonic body wave (P or Sv-wave) with time dependence proportional to $\exp(-i\omega t)$ propagating through the couple stress medium M which we identify as the region $y > 0$, and incident at the interface $y = 0$, with its direction of propagation making an angle θ_0 with the normal to surface. Corresponding to each incident wave, we get waves in couple stress medium M as reflected P and Sv-wave and the SS-wave whose amplitude decays exponentially from the surface and its wave front is at right angles to the surface; and refracted P-and Sv-waves and the SS-wave (traveling with its wave front at right angles to the surface and amplitude decays exponentially from the surface) transmitted to the couple-stress medium M' as shown in fig. 1.

In the region $y > 0$, we write all the variables without a prime and we attach a prime to denote the variables in the region $y < 0$.

The appropriate potentials for this problem are given by

In medium M

$$\phi = A_1 e^{i[\gamma_1 (x \sin \theta_0 - y \cos \theta_0) - \omega t]} + A_2 e^{i[\gamma_1 (x \sin \theta_1 + y \cos \theta_1) - \omega t]} \quad (23)$$

$$\begin{aligned} \psi = B_1 e^{i[\gamma_2 (x \sin \theta_0 - y \cos \theta_0) - \omega t]} + B_2 e^{i[\gamma_2 (x \sin \theta_2 + y \cos \theta_2) - \omega t]} \\ + C_2 e^{[-\gamma y + i(x \gamma_2 \sin \theta_2 - \omega t)]} \end{aligned} \quad (24)$$

where A_1, B_1, A_2, B_2, C_2 are amplitude of various incident and reflected waves.

In medium M'

$$\phi' = A_3 e^{i[\gamma_3 (x \sin \theta_3 - y \cos \theta_3) - \omega t]} \quad (25)$$

$$\psi' = B_3 e^{i[\gamma_4 (x \sin \theta_4 - y \cos \theta_4) - \omega t]} + C_3 e^{[\gamma' y + i(x \gamma_4 \sin \theta_4 - \omega t)]} \quad (26)$$

where A_3, B_3, C_3 are amplitude of various refracted waves. Also $B_1 = 0$ for incident P-wave, $A_1 = 0$ for incident Sv-wave. The Snell's Law is given by

$$\frac{\sin \theta_0}{C_1^*} = \frac{\sin \theta_1}{C_1} = \frac{\sin \theta_2}{C_s} = \frac{\sin \theta_3}{C_1'} = \frac{\sin \theta_4}{C_s'} \quad (27)$$

and

$$\gamma_1 C_1 = \gamma_2 C_s = \gamma_3 C_1' = \gamma_4 C_s' = \omega$$

$$\begin{aligned} C_1^* &= C_1, \text{ for incident P-wave} \\ &= C_s, \text{ for incident Sv-wave} \end{aligned}$$

and

$$C_s^2 = C_2^2 [1 + l^2 \beta_1^2] \quad [\text{Mindlin and Tiersten (1962)}]$$

BOUNDARY CONDITIONS

The boundary conditions at the interface $y = 0$, where the two medium are assumed to be in welded contact are the continuity of tangential displacement, normal displacement, tangential rotational vector, normal force-stress, tangential force-stress and tangential couple-stress, i.e. at $y = 0$.

$$u_x = u'_x, u_y = u'_y, w_z = w'_z, T_{yy} = T'_{yy}, T_{xy} = T'_{xy}, \mu_{yz} = \mu'_{yz} \quad (28)$$

Making use of potentials given by equations (23)-(26) in the boundary conditions (28) and with the help of equations (2),(3),(7) and (26), we obtain the six non-homogeneous equations which can be written as

$$\sum_{j=1}^6 a_{ij} y_j = x_i \quad (i = 1, 2, \dots, 6) \quad (29)$$

where

$$\begin{aligned} a_{11} &= \frac{i\gamma_1}{\gamma^*} \sin \theta_1, & a_{12} &= \frac{i\gamma_2}{\gamma^*} \cos \theta_2, & a_{13} &= \frac{-\gamma}{\gamma^*}, \\ a_{14} &= \frac{-i\gamma_3}{\gamma^*} \sin \theta_3, & a_{15} &= \frac{i\gamma_4}{\gamma^*} \cos \theta_4, & a_{16} &= \frac{-\gamma'}{\gamma^*} \\ a_{21} &= \frac{i\gamma_1}{\gamma^*} \cos \theta_1, & a_{22} &= \frac{-i\gamma_2 \sin \theta_2}{\gamma^*}, & a_{23} &= \frac{-i\gamma_2}{\gamma^*} \sin \theta_2, \\ a_{24} &= \frac{i\gamma_3 \cos \theta_3}{\gamma^*}, & a_{25} &= \frac{i\gamma_4 \sin \theta_4}{\gamma^*}, & a_{26} &= \frac{i\gamma_4 \sin \theta_4}{\gamma^*} \\ a_{32} &= \frac{-\gamma_2^2}{\gamma^{*2}}, & a_{33} &= \frac{(\gamma^2 - \gamma_2^2 \sin^2 \theta_2)}{\gamma^{*2}} \\ a_{35} &= \frac{\gamma_4^2}{\gamma^{*2}}, & a_{36} &= \frac{-1(\gamma'^2 - \gamma_4^2 \sin^2 \theta_4)}{\gamma^{*2}} \\ a_{41} &= \frac{\mu\gamma_1^2}{\gamma^{*2}} (2 \sin^2 \theta_1 - K^2), & a_{42} &= \frac{\mu\gamma_2^2}{\gamma^{*2}} \sin 2\theta_2 \\ a_{43} &= \frac{2i\mu\gamma\gamma_2}{\gamma^{*2}} \sin \theta_2, & a_{44} &= \frac{-\mu'\gamma_3^2}{\gamma^{*2}} (2 \sin^2 \theta_3 - K'^2) \\ a_{45} &= \frac{\mu'\gamma_4^2}{\gamma^{*2}} \sin 2\theta_4, & a_{46} &= \frac{2i\mu'\gamma'\gamma_4}{\gamma^{*2}} \sin \theta_4 \\ a_{51} &= \frac{2\mu\gamma_1^2}{\gamma^{*2}} \sin \theta_1 \cos \theta_1, & a_{52} &= \frac{-\mu}{\gamma^{*2}} \left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2} \right) \end{aligned}$$

$$\begin{aligned}
a_{53} &= \frac{-\mu}{\gamma^{*2}} \left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2} \right), & a_{54} &= \frac{2\mu'\gamma_3^2}{\gamma^{*2}} \sin \theta_3 \cos \theta_3 \\
a_{55} &= \frac{\mu'}{\gamma^{*2}} \left(2\gamma_4^2 \sin^2 \theta_4 - \frac{\omega^2}{C_2'^2} \right) & a_{56} &= \frac{\mu'}{\gamma^{*2}} \left(2\gamma_4^2 \sin^2 \theta_4 - \frac{\omega^2}{C_2'^2} \right) \\
a_{62} &= i\eta \frac{\gamma_2^3}{\gamma^{*3}} \cos \theta_3 & a_{63} &= \frac{\eta}{\gamma^{*3}} (\gamma^3 - \gamma\gamma_2^2 \sin^2 \theta_2) \\
a_{65} &= \frac{i\eta'}{\gamma^{*3}} \gamma_4^3 \cos \theta_4 & a_{66} &= \frac{\eta'}{\gamma^{*3}} (\gamma'^3 - \gamma'\gamma_4^2 \sin^2 \theta_4) \\
a_{31} &= a_{34} = a_{61} = a_{64} = 0 & & (30)
\end{aligned}$$

where

$$\begin{aligned}
\gamma^* &= \gamma_1 \text{ for incident P-wave.} \\
&= \gamma_2 \text{ for incident sv-wave.} & (31)
\end{aligned}$$

(a) For incident P-wave.

$$x_1 = -a_{11}, x_2 = a_{21}, x_3 = a_{31}, x_4 = -a_{41}, x_5 = a_{51}, x_6 = a_{61}.$$

(b) For incident Sv-wave.

$$x_1 = a_{12}, x_2 = -a_{22}, x_3 = -a_{32}, x_4 = a_{42}, x_5 = -a_{52}, x_6 = a_{62}$$

The amplitude ratios of reflected and transmitted wave are,

$$y_1 = \frac{A_2}{A^*}, y_2 = \frac{B_2}{A^*}, y_3 = \frac{C_2}{A^*}, y_4 = \frac{A_3}{A^*}, y_5 = \frac{B_3}{A^*}, y_6 = \frac{C_3}{A^*}$$

where

$$\begin{aligned}
A^* &= A_1, \text{ for incident P-wave} \\
&= B_1, \text{ for incident Sv-wave.} & (32)
\end{aligned}$$

SPECIAL CASES

CASE I. If all the elastic moduli in medium M' vanish, then the problem reduces to reflection of plane waves at a flat free boundary $y = 0$, in this case, the boundary conditions reduce to

$$T_{yy} = T_{xy} = \mu_{yz} = 0.$$

Therefore, from equation (29), we obtain

(a) For incident P-Wave

$$\frac{A_2}{A_1} = \left[\gamma\gamma_1^2 \gamma_2^2 \sin 2\theta_1 \sin 2\theta_2 (\gamma^2 + \gamma_2^2 \cos^2 \theta_2) \right]$$

$$-\left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2}\right)\left(2\gamma_1^2 \sin^2 \theta_1 - \frac{\omega^2}{C_2^2}\right)\left(\gamma^3 - \gamma\gamma_2^2 \sin^2 \theta_2 - 1\gamma_2^3 \cos \theta_2\right)\Bigg]/D_1$$

$$\frac{B_2}{A_1} = \left[-2\gamma_1^2 \sin 2\theta_1 (\gamma^3 - \gamma\gamma_2^2 \sin^2 \theta_2) \left(2\gamma_1^2 \sin^2 \theta_1 - \frac{\omega^2}{C_2^2}\right) \right] / D_1$$

$$\frac{C_2}{A_1} = \left[-2i\gamma^2 \gamma_2^3 \cos \theta_2 \sin 2\theta_1 \left(2\gamma_1^2 \sin^2 \theta_1 - \frac{\omega^2}{C_2^2}\right) \right] / D_1$$

where

$$D_1 = \gamma\gamma_1^2\gamma_2^2 \sin 2\theta_1 \sin 2\theta_2 (\gamma^2 + \gamma_2^2 \cos^2 \theta_2)$$

$$+ \left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2}\right) \left(2\gamma_1^2 \sin^2 \theta_1 - \frac{\omega^2}{C_2^2}\right)$$

$$(\gamma^3 - \gamma\gamma_2^2 \sin^2 \theta_2 - 1\gamma_2^3 \cos \theta_2)$$

(b) For incident Sv-wave

$$\frac{A_2}{B_1} = \left[(2\gamma_2^2 \sin 2\theta_2) \left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2}\right) (\gamma\gamma_2^2 \cos^2 \theta_2 + \gamma^3) \right] / D_2$$

$$\frac{B_2}{B_1} = \left[(\gamma_1^2 \sin 2\theta_1) (\gamma_2^2 \sin 2\theta_2) (\gamma\gamma_2^2 \cos^2 \theta_2 + \gamma^3) \right]$$

$$-\left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2}\right) \left(2\gamma_1^2 \sin^2 \theta_1 - \frac{\omega^2}{C_2^2}\right) (\gamma^3 - \gamma\gamma_2^2 \sin^2 \theta_2 + i\gamma_2^3 \cos \theta_2) \Bigg] / D_2$$

$$\frac{C_2}{B_1} = \left[2i\gamma_2^3 \cos \theta_2 \left(2\gamma_1^2 \sin^2 \theta_1 - \frac{\omega^2}{C_2^2}\right) \left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2}\right) \right] / D_2$$

where

$$D_2 = (\gamma_1^2 \sin 2\theta_1) (\gamma_2^2 \sin 2\theta_2) (\gamma\gamma_2^2 \cos^2 \theta_2 + \gamma^3)$$

$$+ \left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2}\right) \left(2\gamma_1^2 \sin^2 \theta_1 - \frac{\omega^2}{C_2^2}\right) (\gamma^3 - \gamma\gamma_2^2 \sin^2 \theta_2 + i\gamma_2^3 \cos \theta_2)$$

The above amplitude ratios are same as obtained by Graff and Pao [3].

CASE II. If we let $\eta = \eta' = 0$, then both media are reduced to the isotropic elastic half spaces in welded contact and we obtain a set of four non-homogeneous equations as

$$\sum_{j=1}^4 C_{ij} y_j = x_i \quad (i = 1, 2, 3, 4) \quad (33)$$

$$C_{11} = \frac{i\gamma_1}{\gamma^*} \sin \theta_1, \quad C_{12} = \frac{i\gamma_2}{\gamma^*} \cos \theta_2, \quad C_{13} = \frac{-i\gamma_3}{\gamma^*} \sin \theta_3$$

$$C_{14} = \frac{i\gamma_4}{\gamma^*} \cos \theta_4, \quad C_{21} = \frac{i\gamma_1}{\gamma^*} \cos \theta_1, \quad C_{22} = \frac{-i\gamma_2 \sin \theta_2}{\gamma^*}$$

$$C_{23} = \frac{i\gamma_3}{\gamma^*} \cos \theta_3, \quad C_{24} = \frac{i\gamma_4 \sin \theta_4}{\gamma^*}$$

$$C_{31} = \frac{\mu\gamma_1^2}{\gamma^{*2}} (2\sin^2 \theta_1 - K_1^2), \quad C_{32} = \frac{\mu\gamma_2^2}{\gamma^{*2}} \sin 2\theta_2$$

$$C_{33} = \frac{-\mu'\gamma_3^2}{\gamma^{*2}} (2\sin^2 \theta_3 - K_1'^2), \quad C_{34} = \frac{\mu'\gamma_4^2}{\gamma^{*2}} \sin 2\theta_4$$

$$C_{41} = \frac{2\mu\gamma_1^2}{\gamma^{*2}} \sin \theta_1 \cos \theta_1, \quad C_{42} = \frac{-\mu}{\gamma^{*2}} \left(2\gamma_2^2 \sin^2 \theta_2 - \frac{\omega^2}{C_2^2} \right)$$

$$C_{43} = \frac{2\mu'\gamma_3^2}{\gamma^{*2}} \sin \theta_3 \cos \theta_3, \quad C_{44} = \frac{\mu'}{\gamma^{*2}} \left(2\gamma_4^2 \sin^2 \theta_4 - \frac{\omega^2}{C_2'^2} \right)$$

(c) For incident p-wave.

$$x_1 = -C_{11}, \quad x_2 = C_{21}, \quad x_3 = -C_{31}, \quad x_4 = C_{41}$$

(d) For incident Sv-wave.

$$x_1 = C_{12}, \quad x_2 = -C_{22}, \quad x_3 = C_{32}, \quad x_4 = -C_{42}$$

and

$$y_1 = \frac{A_2}{A^*}, \quad y_2 = \frac{B_2}{A^*}, \quad y_3 = \frac{A_3}{A^*}, \quad y_4 = \frac{B_3}{A^*}$$

where γ^* and A^* and given by equations (31) and (32) respectively.

The above results are same as obtained in Ewing, Jardetzky and Press [2].

NUMERICAL DISCUSSION

To discuss the problem numerically when P-wave and Sv-wave are incident, we consider the following values of relevant parameters for couple-stress elastic solid.

For medium M ; ($y > 0$)

$$\lambda = 2.2. \text{ dyne/cm}^2, \quad \mu = 0.81 \text{ dyne/cm}^2, \quad \rho = 2.6 \text{ gm/cm}^3$$

For medium M' ($y < 0$)

$$\lambda' = 0.96 \text{ dyne/cm}^2, \quad \mu' = 0.71 \text{ dyne/cm}^2, \quad \rho' = 1.93 \text{ gm/cm}^3$$

The non-dimensional constants are $l^2 \gamma_2^2 = 0.1$, $l^2 \gamma_4^2 = 0.01$ and $\eta'/\eta = 0.05$.

For the above values of relevant constants, the system of equations (29) is solved for amplitude ratios by using Gauss elimination method for different angle of incidence varying from 0° to 90° . The variations of these amplitude ratios with the angle of incidence for incident P-wave and Sv-wave have been shown graphically in Fig. (2)-(19). The solid lines in these figures correspond to the variation of amplitude ratios in couple-stress elastic medium, whereas the dashed lines correspond to the variation of amplitude ratios for elastic medium.

CASE I. Couple-stress elastic and couple-stress elastic interface

Subcase (a) : Incident P-wave.

The comparison of amplitude ratio $|z_1|$ of reflected P-wave in couple stress elastic medium, with amplitude ratio $|z_1|$ of reflected P-wave in elastic medium shows that the effect of couple-stress increases the value of amplitude ratio $|z_1|$ of elastic medium with increase in angle of incidence. There is a sharp difference between the values of these amplitude ratios initially, but as the angle of incidence increases the difference decreases slowly and at $\theta_0 = 90^\circ$ the both amplitude ratios approaches the same value, these variation are shown in figure(2).

Fig. 3 depicts the variations of the reflection coefficients $|z_2|$ of the reflected Sv-wave in couple stress elastic and in elastic medium. The amplitude ratio $|z_2|$ of reflected Sv-wave in couple-stress elastic medium increases more rapidly than the amplitude ratio $|z_2|$ of reflected Sv-wave in elastic medium however, there is a sharp difference between the max. value of both amplitude ratios.

When we compare the amplitude ratio $|z_4|$ of refracted P-wave in couple-stress medium with amplitude ratio $|z_4|$ of refracted P-wave in elastic medium, we observe that both amplitude ratios decrease monotonically and approaches to same value at $\theta_0 = 90^\circ$, though the initial value of the amplitude ratio $|z_4|$ of couple-stress medium is more than the amplitude ratio $|z_4|$ of elastic medium due to couple-stress effect. All these variations are shown in Fig. 4. In Fig. 5, the amplitude ratio $|z_5|$ of refracted Sv-wave in elastic medium has been magnified by multiplying its original value by 10. The comparison of amplitude ratio $|z_5|$ of refracted Sv-wave in couple-stress medium with amplitude ratio $|z_5|$ reveals that both approaches to nearby same value at $\theta_0 = 90^\circ$, but has a sharp difference at the initial value. The variation of the amplitude ratios of $|z_3|$ and $|z_6|$ of reflected ss-wav have been depicted in Figs.6 and 7 respectively.

Sub Case (b) : Incident SV-wave

The amplitude ratio $|z_1|$ of reflected P-wave in couple stress elastic medium decreases sharply up to $\theta_0 = 10^\circ$, decreases monotonically for $10^\circ \leq \theta_0 \leq 33^\circ$, oscillates for the range $33^\circ \leq \theta_0 \leq 72^\circ$ and again decreases as θ_0 increases further. The value of the amplitude ratio $|z_1|$ of reflected P-wave in elastic medium has been magnified by multiplying its original value by 10 and these variations are shown in Fig. 8.

Fit 9 depicts the variations of the amplitude ratio $|z_2|$ of the reflected Sv-wave in couple stress and elastic medium and it is observed that there is sharp difference between the maximum values of these amplitude ratios.

The amplitude ratio $|z_3|$ of refracted P-wave in elastic medium has its max. value at $\theta_0 = 2^\circ$ and decreases first sharply and then gradually. The value of the amplitude ratio $|z_4|$ of refracted P-wave in couple stress medium has been magnified by multiplying its original value by 10 and these variations are shown in Fig. 10.

Fig. 11 shows the variations of the amplitude ratio $|z_5|$ of reflected Sv-wave in couple stress elastic medium and the amplitude ratio $|z_5|$ of refracted Sv-wave in elastic medium.

The variations of the amplitude ratios $|z_3|$ and $|z_6|$ of the reflected SS-wave and refracted SS-wave respectively in couple stress elastic medium have been shown in Figs. 12 and 13 respectively.

CASE II. Couple-Stress elastic half-space

Subcase (a) : Incident P-wave

When we compare the amplitude ratio $|z_1|$ of reflected P-wave in couple-stress elastic half space with the amplitude ratio $|z_1|$ of reflected P-wave in elastic half space, we observe that due to the effect of couple stress value of the amplitude $|z_1|$ of elastic half space increases, though both of these approaches same value at $\theta_0 = 90^\circ$ and these variations are shown in Fig. 14.

The amplitude ratios $|z_2|$ of reflected Sv-wave in coupled-stress elastic half space and the amplitude ratio $|z_2|$ of reflected Sv-wave in elastic half space have almost same value except the max. value and around it and these variations are shown in Fig. 15.

The variation of amplitude ratio $|z_3|$ of SS-wave in coupled half space has been shown in Fig. 18.

Subcase (b) : Incident Sv-wave

The comparison of the amplitude ratio $|z_1|$ of reflected P-wave in coupled stress elastic half space with the amplitude ratio $|z_1|$ of reflected P-wave in elastic half space has been shown in fig. 16. The value of amplitude ratio $|z_1|$ of elastic half space has been magnified by multiplying its original value by 10, which is almost similar to the max. value of $|z_1|$ of elastic half space.

The amplitude ratio $|z_2|$ of reflected Sv-wave in coupled half space has its max. value at $\theta_0 = 49^\circ$, which is very large in comparison to the max. value of the amplitude ratio $|z_2|$ of reflected Sv-wave in elastic half-space and these variations are shown in Fig. 17.

The variation of amplitude ratio $|z_3|$ of SS-wave in coupled half space has been shown in Fig 19.

CONCLUSION

The analytical expressions for reflection and transmission coefficients of various reflected waves are derived. The variations of the reflection and transmission coefficients of various reflected waves have been depicted graphically. Some particular cases have been deduced. It may be concluded that, the effect of couple stress plays an important role in a reflection and refraction phenomena. The model adopted in this paper is one of the realistic forms of the earth model and it may be of interest for experimental seismologists.

LIST OF SYMBOLS

- λ, μ = Lamé's constants
 η = The bending twist modulus,
 ρ = The density
 ϵ_{ijk} = Alternating Tensor
 δ, ω_j = Kronecker Delta, rotational vector respectively
 η' = The bending of twisting modulus
 l = The dimension of length and it carries the effect of couple-stress

REFERENCES

1. Aggarwal, H.R. and Alverson, A.C. (1969), 'Effects of couple-stresses on the diffraction of plane elastic waves by cylindrical discontinuities', Int J. solids structure 5, 491-511.
2. Ewing, W.M., Jardetzky, W.S., and Press, F. (1927), 'Elastic waves in layered media', McGraw-Hill, New York.
3. Graff, K.F. and Pao, Y.H. (1967), 'The effects of couple-stresses on the propagation and reflection of plane waves in an elastic half space', J. sound vib., 6(2), 217-229.

4. Love, A.E.H. (1927), 'Mathematical theory of elasticity', Dover Publication, fourth edition, New York.
5. Mindlin, R.D. and Tiersten, H.F. (1962), 'Effects of couple-stresses in linear elasticity', Arch. Rat. Mech. Analysis, 11, 415-418.
6. Sengupta, P.R. and Ghosh, B.C. (1974), 'Effect of couple stresses on the surface waves in elastic media,' Gerlands Beitr. Geophysics, 83(4); 309-318.
7. Sengupta, P.R. and Benerji, D.K. (1978), 'Effects of couple-stresses on propagation of waves in an elastic layer immersed in an infinite liquid', Int. J. of pure and Appl. Maths, 9, 17-28.
8. Kumar and Singh (1998), Reflection and refractions of planes waves at an interface between micropolar elastic solid and viscoelastic solid. Int. J. Engg. Sci., 36, 119-135.
9. Kumar and Singh (2000), Reflection of plane waves at a planer viscoelastic micropolar interface. Indian J. Pure and Appl. Math., 31,287-303.
10. Singh, B. and Kumar, R. (2004), Studied the problem of wave propagation in a micropolar elastic solid with stretch, Proc.Nat.Acad.Sci.India.74,123-133

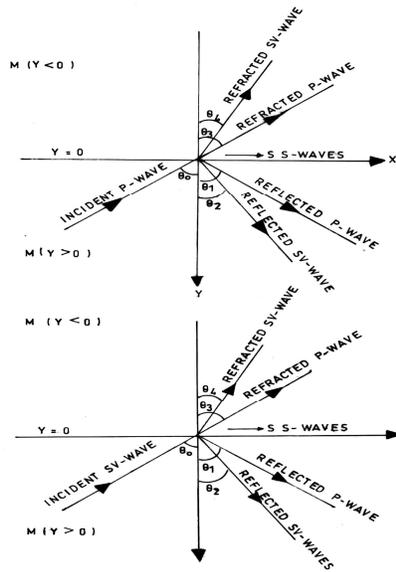
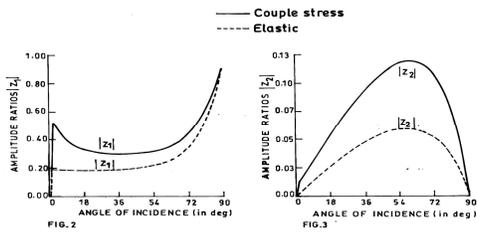
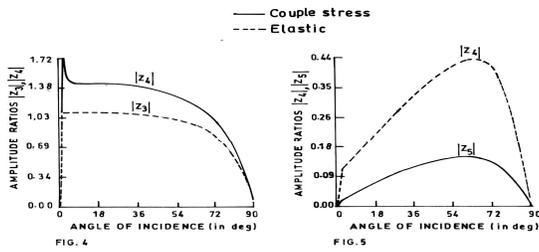


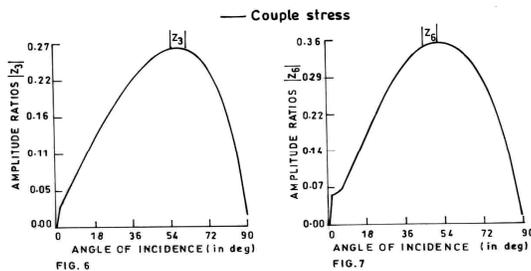
Fig.1 Geometry of the problem.



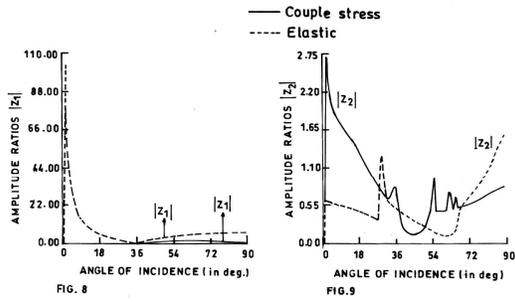
Figs 2-3 Variations of amplitude ratios with incident angle of P-wave at couple stress/couple-stress elastic interface.



Figs 4-5 Variations of amplitude ratios with incident angle of P-wave at couple stress/couple-stress elastic interface



Figs 6-7 Variations of amplitude ratios with incident angle of P-wave at couple stress/couple-stress elastic interface



Figs 8-9 Variations of amplitude ratios with incident angle of Sv-wave at couple stress/couple stress-elastic interface

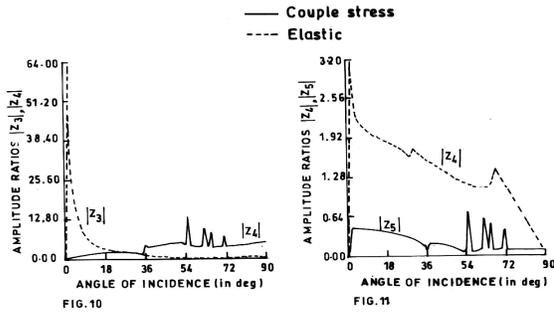
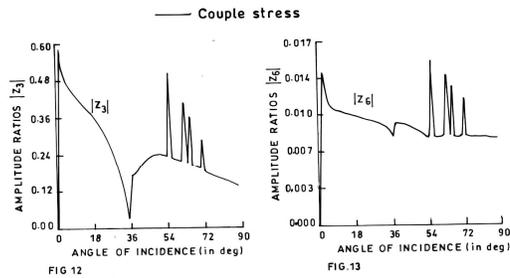
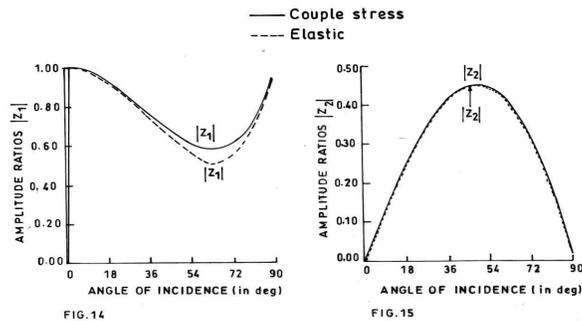


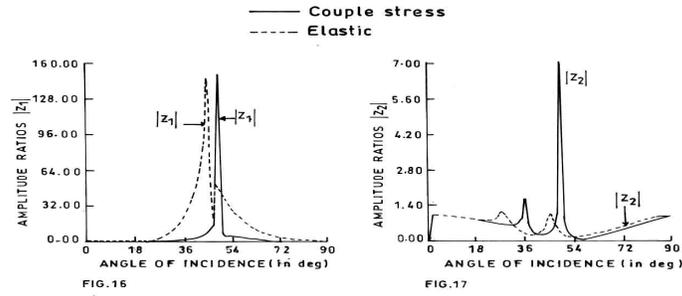
Fig. 10-11 Variations of amplitude ratios with incident angle of Sv-wave at couple stress/couple stress-elastic interface



Figs 12-13 Variations of amplitude ratios with incident angle of Sv-wave at couple stress/couple stress-elastic interface



Figs 14-15 Variations of amplitude ratios with incident angle of P-wave at couple stress elastic half-space.



FIGS 16-17 Variations of amplitude ratios with incident angle of Sv-wave at couple stress elastic half-space.

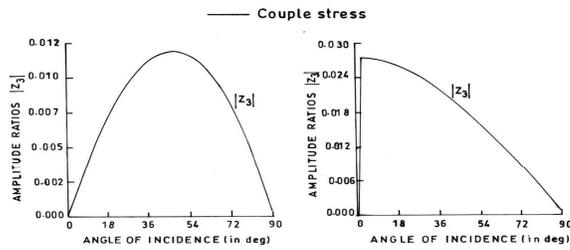


Fig.18 Variations of amplitude ratios with incident angle of P-wave at couple stress elastic half-space.

Fig.19 Variations of amplitude ratios with incident angle of Sv-wave at couple stress elastic half-space.

An Integrated Multistage And Multicriteria Analytical Hierarchy Process GIS Model For Landfill Siting: A Model

SS.Asadi ¹,B.V.T.Vasantha Rao ² M.V.Raju ³and V.Padmaja ⁴

¹Associte professor,Dept.of. Civil Engineering, KL University, Green fields, Vaddeswaram-522502,Guntur(D.t),A.P,India

²Assistant Professor,Dept.of .Civil Engineering,p.v.p.siddhardha Institute of Technology,Kannure, Vijayawada.

³ Assistant Professor,Dept.of.Civil Engineering, Vignan University, Vadllamudi, Guntur(D.t).

⁴Centre for Environment, J.N.T.University, kukatpally, Hyderabad

Abstract: Human activity, either domestic or industrial, produces waste materials to greater or lesser extent always. Each city produces tons of solid wastes daily from household, hospitals, industry, offices and market centers. By volume, municipal solid waste is by far the largest component of overall urban waste stream and includes a variety of potentially toxic substances making its disposal a problem. However, either due to resource crunch or inefficient infrastructure, not all of this waste gets collected and transported efficiently to final dumpsites leading to an improper management and disposal, resulting in serious impacts on health and problems to surrounding environment. The efficacy of solid waste disposal depends upon selection of proper site and there are several issues that have impact for site selection. Site selection is therefore one of the most critical areas of municipal planning involving a multi-disciplinary approach and a broad spectrum of considerations. As such, it is a multi-criteria decision-making process, a much-needed system for proper municipal solid waste management. This paper attempts to demonstrate the capabilities and utility of remote sensing and GIS technology for selection of suitable sites for waste disposal for Hyderabad City, India. The study illustrates the application of GIS techniques in the field of solid waste management, through the typical problem of preliminary disposal site selection, using a multi-criteria decision making technique called the Analytic Hierarchy Process (AHP) which provides a systematic approach for assessing and integrating the impacts of various factors, involving several levels of, dependent and independent, qualitative and quantitative information. Remote sensing images and

Survey of India topomaps are used to extract information on wastelands and other land use features, geology, hydrogeomorphology, drainage, road network and slope of the area. The approach used in this project was firstly to exclude all areas where the development of a waste disposal site would not be permitted viz., close proximity to residential areas, airfields, mountainous areas, nature reserves, indigenous forests, geological faults, the coast, dams or rivers. Once these areas had been identified, the remaining areas were then rated according to the geological, hydrological, topographical and environmental characteristics. From the combination of these factors, favorable areas were identified. In GIS, all conceivable requirements (e.g. site should be at least 500 meters from the nearest dwelling unit) are initially specified, spatial data integrated and overlaid and based on the final output obtained suitable environmentally benign sites for waste disposal are identified.

Keywords: Solid waste, Multi criteria decision-making, Analytical hierarchy Process, Remote sensing and GIS

Introduction

In many countries with increase in population and the rising demand for food and other essentials, there has been a rise in the amount of solid waste being generated making its management and disposal problematic. The accumulation and improper disposal of waste leads to environmental pollution and accelerates the spread of communicable diseases (George Tchobanoglous, 1993). One of the serious and growing potential problems in most large urban areas is the shortage of land for waste disposal, and Hyderabad city in India with an alarming pace of population explosion and urbanization is no exception. Management of municipal solid waste (MSW) in urban centers is becoming more complex due to scarcity of land for disposal (Baburani devi and Bhojar, 2003). The waste generators find it difficult to dispose their wastes without causing environmental disturbance, as very few appropriate disposal facilities are available. An appropriate landfill site should therefore be selected carefully by considering both regulations and constraints on other sources in order to prevent environmental, economic and ecological impacts.

Keeping this in view, the present study is carried out with an objective of identifying a suitable site for disposal of municipal solid waste generated in Hyderabad city using Geographic Information System (GIS) and Analytical Hierarchy Process (AHP). According to the existing records of the Municipal Corporation of Hyderabad (MCH), the total solid waste generated in Hyderabad city is about 2200 MT/day of which 1500MT is disposed by landfilling and the remaining 700MT is utilized for power generation (JNTU, 2002). With the Autonagar dumping site being closed in 2005, there exists only one site operating at present located at Jawaharnagar for disposing this waste. Keeping in view the need for disposal sites, an attempt has been made in this study to identify potential sites for disposal of solid waste generated in Hyderabad city.

Objectives Of The Study

- To study the status of existing scenario of solid waste management in Hyderabad city by analyzing the trends in population growth, waste generation, existing transfer stations and solid waste disposal sites in order to evolve a strategic plan for future years.
- To create spatial digital database comprising base map, land use/ land cover, soil, slope, physiography, geology, geomorphology, drainage pattern, ground water potential, infiltration rate, watershed, transportation network and wasteland map with the help of IRS –1D, PAN & LISS-III merged satellite imagery and Survey of India (SOI) toposheets along with ground truth analysis on ARC/INFO GIS platform.
- To evaluate AHP and assign weightage factors for different digital thematic layers and to integrate them in GIS for evaluating risk suitability and to find probable solid waste disposal sites in and around Hyderabad based on this integrated study.

Study Area

Hyderabad city is situated in the Krishna basin and the river Musi, which is a tributary of river Krishna, passes through the city and bifurcates it into Northern and Southern Hyderabad. The study region covers an area of 179Km² and is situated between 78°22'30" & 78°32'30" east longitude & between 17°18'30" & 17°28'30" north latitude. The ground levels vary from 487 meters to 610 meters above mean sea level. The region of interest for site selection includes all area, which falls within the buffer distance of 50km from the center of Hyderabad city. This area

comprises of Hyderabad Urban Development Area, parts of Rangareddy, Nalgonda, Medak and Mahabubnagar districts of Andhra Pradesh. It is covered by toposheet No. 56K on 1:2,50,000 scale. The city stands on gray and pink granites as foundation materials, which is suitable for building construction. According to 2001 census Hyderabad is one of the largest metropolis of India with a population of 38,29,753. The population of Hyderabad has increased from 0.448 million in 1901 to 1.429 million in 1961, between 1981 to 1991 the population went up to 4.34 million and the rate of the growth so far is 67.04% (Handbook of Statistics, Hyderabad District, 2001).

The physical composition of solid waste generated in Hyderabad includes 4% paper, <1% glass, 3.5% plastics, 40-45% biodegradable matter, 45-50% inert waste/silt/debris and 2-3% others. The average moisture content of the garbage is 21.05% where as the dump yard it is 19.25%. The net calorific value of Hyderabad household garbage is 1223 cal/gm (ASCI, 2004).

Methodology

The site selection is carried out in two phases. In the first phase, criteria to be considered for municipal solid waste disposal site selection are identified and broadly grouped into exclusionary and non-exclusionary criteria. These criteria are defined using the standards given in Municipal Solid Waste (Management and Handling) Rules 2000 and includes landfill area required, proximity to lakes and rivers, proximity to highways, distance to population centers, slope, etc. The non-exclusionary criteria are further categorized into four categories viz., geological criteria consisting of geology and soil characteristics, topographical criteria which include the land use, geomorphology and slope patterns, hydrological criteria comprising of groundwater level, potential and infiltration characteristics and environmental criteria encompassing groundwater and air quality (Figure 1).

Exclusionary Criteria:

- No landfill should be constructed within 200mts of any lake or pond
- No landfill must be constructed within 100mts of navigable river or stream
- No landfill should be constructed within a 100 yr flood plain
- No landfill should be constructed within 200mts of any state or national highway

- A landfill site must be at least 1000mts from a notified habitat area
- No landfill should be constructed within 300mts of public parks
- No landfill should be constructed within critical habitat areas
- No landfill should be constructed within wetlands
- A landfill should not be constructed in areas where water table is less than 5m below ground surface
- No landfill should be constructed within 2000mts of an airport
- No landfill should be constructed within 500mt of any water supply well
- A landfill should not be sited in a coastal regulation zone
- A landfill should not be located in potentially unstable zones such as landslide prone areas, fault zone etc

Non-Exclusionary Criteria:

Geological criteria - Geology, Structures and Soil

Topographical criteria - Road network, Slope, Land use/ Land cover, Geomorphology

Environmental criteria - Groundwater quality and Air quality

Hydrological criteria - Groundwater potential, Groundwater infiltration and Groundwater table

Once the parameters are organized into a decision hierarchy, Relative Importance Weightage (RIW) of each parameter over the other is calculated by pair-wise comparison using the 9-point scale shown in Table 1. The RIWs are the normalized eigenvectors corresponding to the maximum eigen values of the pair-wise comparison matrices constructed at each level of the decision hierarchy. The RIW assigned to each hierarchy element is determined by normalizing the eigenvector of the decision matrix. Eigenvector values are estimated by multiplying all the elements in a row and taking the nth root of the product, where n is the number of row elements (Saaty, 1980 and Siddiqui et al, 1996). Normalization of the eigenvector is accomplished by dividing each eigenvector element by the sum of the eigenvector elements of the decision matrix. Once the RIW of each element of each theme or layer to be considered for site selection are calculated, the individual weightages are aggregated/summarized to obtain a final suitability index (Ertan Yesilnacar and Vedat Doyuran, 2000 and Aditya et al, ICORG).

In the second phase, ERDAS and ARC/INFO software are used to generate topographic, thematic and spatial data corresponding to layers including settlements, roads, topography, geology, land use/ land cover, geomorphology, aquifers and surface water, soil etc. from fused data of IRS-1D PAN and LISS-III satellite imagery, Survey of India topomaps, existing datasets and field data. The methodology adopted for the study is shown in Figure 2. The land use/ land cover map prepared using visual interpretation technique from fused satellite imagery is presented in Figure 3. ArcView GIS tools were used to create buffers, calculate distances, and screen out areas that failed any minimum criteria (such as minimum distance from airports and bodies of water). After screening out unacceptable areas within each theme corresponding to some criteria, all themes are overlaid to prepare a composite suitability map. Areas not satisfying the minimum landfill criteria were eliminated and other parcel groups that remained are further evaluated based on suitability indices values and classified into excellent, good, moderate, poor and very poor sites. The sites within each of these five classes are further evaluated for their attributes such as distance from the point of waste generation, population density surrounding the site, proximity to settlement areas, presence of scrub forest, total area covered etc. resulting in selection of five best suitable sites within excellent, good and moderate suitability class for the entire study area.

Results And Discussions

By applying AHP and GIS various layers are categorized according to Relative Importance Weightage (RIW). For final weight all individual weightage are summarized and evaluated to obtain a final suitability index. As each of the RIW includes its importance from the first level, the final cumulative values will give more accurate results. According to the weightage allotted and the suitability index value obtained, entire area is categorized into five classes as excellent class with suitability index ranging from 0.5 to 0.6, good class with suitability index ranging from 0.4 to 0.5, moderate class with suitability index ranging from 0.3 to 0.4, poor class with suitability index ranging from 0.2 to 0.3 and very poor class with suitability index ranging from 0.1 to 0.2 with respect to landfill siting. Higher the suitability index, the more suited is the site for waste disposal and lower the value, lower is the suitability. The suitability map prepared for the present study is depicted in Figure 4. After the GIS analysis led to a short list of sites, the attribute evaluation (distance from the point of waste generation, area covered, distance to

nearest road or water body, population density surrounding the site etc.) of each of these individual sites was performed to determine which site possessed the best compromise of features for developing a landfill and are ranked accordingly. Of the many sites identified, sites selected for solid waste dumping based on attribute evaluation are presented in Table 2.

Conclusions

The study presents an elaborate introduction to GIS, along with the operational details of building a database in a GIS environment. This study also illustrates, how the capabilities of GIS can be utilized for addressing problems of spatial nature, with the efficiency of handling large volumes of spatial data, and with speed, which is quite typical of computer-based operations. The choice of the decision factors for preliminary site selection for solid waste disposal was guided predominantly by the availability of data. In the process of the site selection study, a comprehensive description of AHP has been provided along with the records of the application of AHP for landfill siting. The present study is an attempt, in a similar direction, to preliminary site selection for open dumping of solid wastes of Hyderabad city. The gap between the preliminary and the final stage of site selection can be laid down as the scope for future work. The accuracy of the final modelling, integrating the GIS database with the AHP, depends on the quality of the database and on the fundamental assumptions of the AHP technique. The construction of the decision matrices, largely determines the final output, but the construction of these matrices are essentially subjective, as any other decision-making process, based on judgment.

References

1. George Tchobanoglous, Hilary Theisen, and Samuel Vigil, Integrated Solid Waste Management-Engineering Principles And Management Issues, Mc. Graw Hill, New York, 1993.
2. Babyrani Devi, S. and Bhoyar, R.V., Feasibility Of Some Treatments For Improving The Composting Of Municipal Solid Waste, Indian Journal of Environmental Health, Vol. 45, No. 3, pp 231-234, July 2003.
3. Inventorization of Municipal Solid Waste Management Practices in Municipalities/ Corporations of Andhra Pradesh, JNTU Project Report (Current Status On Management Of Municipal Solid Waste Hyderabad District), 2002.
4. District Census Handbook of Hyderabad, Director of Census Operations, Andhra Pradesh, Census of India, 2001.
5. Brainstorming Meeting on Improving the Solid Waste Disposal System in Hyderabad, Background Papers, Administrative Staff College of India (ASCI), Bella Vista: Hyderabad, October 2004.
6. A Support Manual For Municipal Solid Wastes (Management and Handling) Rules, 2000, Central Pollution Control Board (CPCB), Ministry of Environment and Forests, New Delhi.
7. Saaty, T. L., The Analytic Hierarchy Process, McGraw-Hill Book Co., New York, 1980.
8. Siddiqui, M.Z., Everett, J.W., and Vieux, B.E., Landfill Siting Using Geographic Information Systems: A Demonstration, Journal of Environmental Engineering, ASCE, 122 (6), pp 515-523, 1996.
9. Ertan Yesilnacar and Vedat Doyuran, Selection Of Settlement Areas Using GIS And Statistical Method (Spatial-AHP), Project Expo 2000, International Forum for Graduates and Young Researchers at EXPO 2000, Shaping the Future Hannover, The World Exposition in Germany, 2000.
10. Aditya Agrawal, Manish Katyan, Onkar Dikshit and Rajiv Sinha, Preliminary Selection Of Solid Waste Disposal Sites Using Remote Sensing And GIS Techniques, Proceedings of International Conference on Remote Sensing and GIS (ICORG), Vol. 1, pp 185-190, 2001.

Table 1: Analytic Hierarchy Measurement Scale

| Reciprocal Measure of Intensity of Importance | Definition | Explanation |
|--|---|--|
| 1 | Equal importance | Two activities contribute equally to the objective |
| 3 | Weak importance of one over another | Experience and judgment slightly favour one activity over another |
| 5 | Essential or strong importance | Experience and judgment strongly favour one activity over another |
| 7 | Demonstrated importance | An activity is strongly favoured and its dominance is demonstrated in practice |
| 9 | Absolute importance | The evidence favouring one activity over another is of the highest possible order of affirmation |
| 2,4,6,8 | Intermediate values between two adjacent judgments | When compromise is needed |
| Reciprocal of the above | If activity I has one of the above non-zero numbers assigned to it when compared with activity j, then j has the reciprocal value when compared with i. | |

Table 2 Suitable sites for solid waste dumping

| Excellent | Location |
|------------------|--|
| Site 1 | Near Kuntlur and Annaram villages – east of Hyderabad |
| Site 2 | Between Pocharam and Yemnapet villages – east of Hyderabad |
| Site 3 | Near Lakdaram – NW of Hyderabad |
| Site 4 | Near Kisara, Peddaparvatapuram and Bhogawaram - NE |
| Site 5 | Near Pratapasingaram, Koremalla and Choudariguda – E of Hyderabad |
| Good | Location |
| Site 1 | Between Nerapalli, Polkampalli and Manyaguda villages – SE of Hyderabad |
| Site 2 | Between Seriguda and Turka Yemjal - SE of Hyderabad |
| Site 3 | Between Narapalli and Kachwani Singaram – east of Hyderabad |
| Site 4 | Between Ismailkhanguda and Pocharam – SE of Hyderabad |
| Site 5 | Near Upparpalli and Tumukunta – SE of Hyderabad |
| Moderate | Location |
| Site 1 | Near Yadagiripalli and Kisara – east of Hyderabad |
| Site 2 | Near Madhawaram, Gandigudem and Kazipalli – NW of Hyderabad |
| Site 3 | Between Kondapuram, Charlapalli and Ghatkesar villages–east of Hyderabad |
| Site 4 | Between Srirangaram, Dablipur and Girmapuram – north of Hyderabad |
| Site 5 | Between Kisara, Bhogawaram and Peddaparvatapuram-east of Hyderabad |

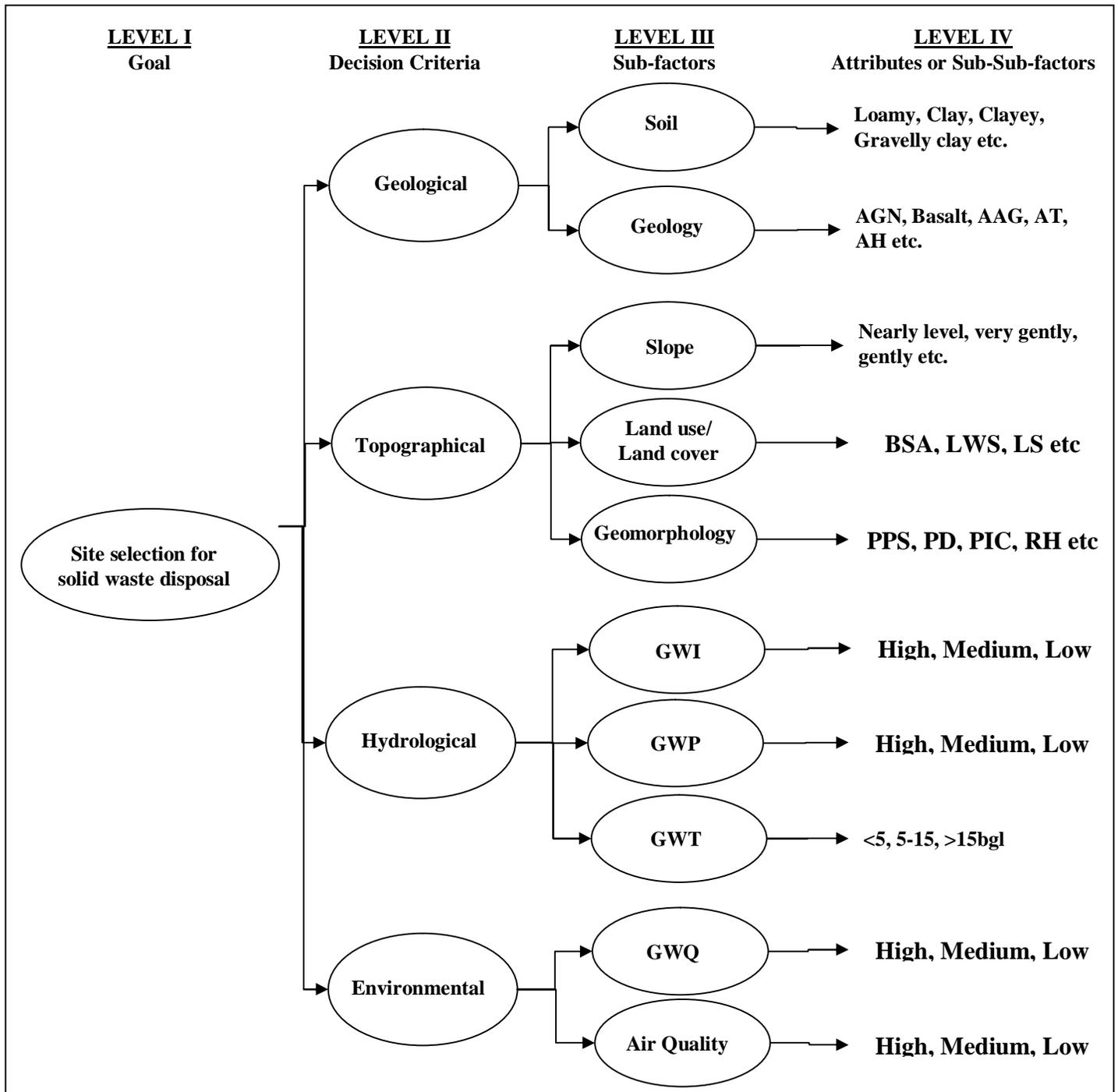


Figure 1: Structuring of decision hierarchy for a typical solid waste disposal site selection

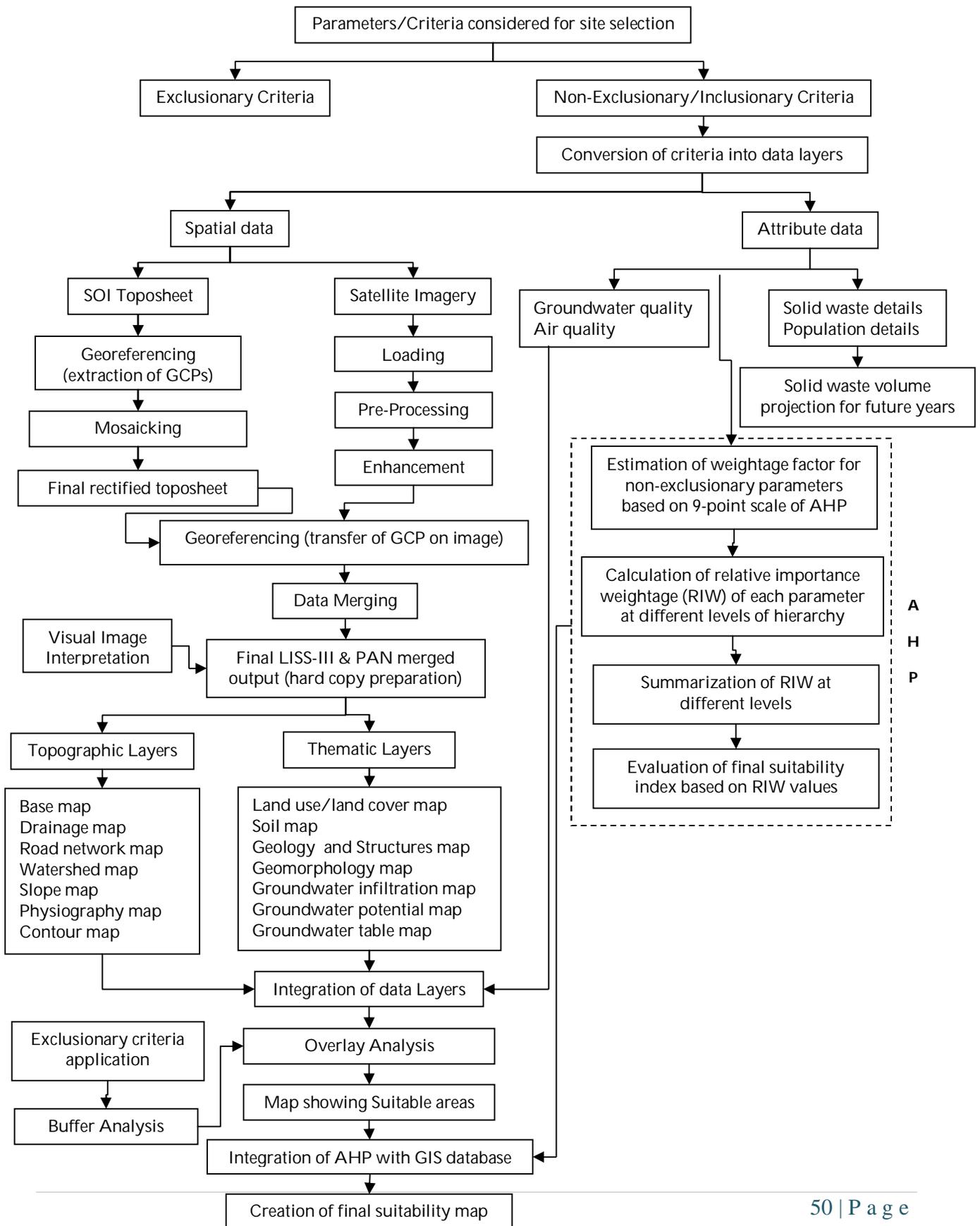


Figure 2: Flow chart showing the overall methodology adopted for the present study

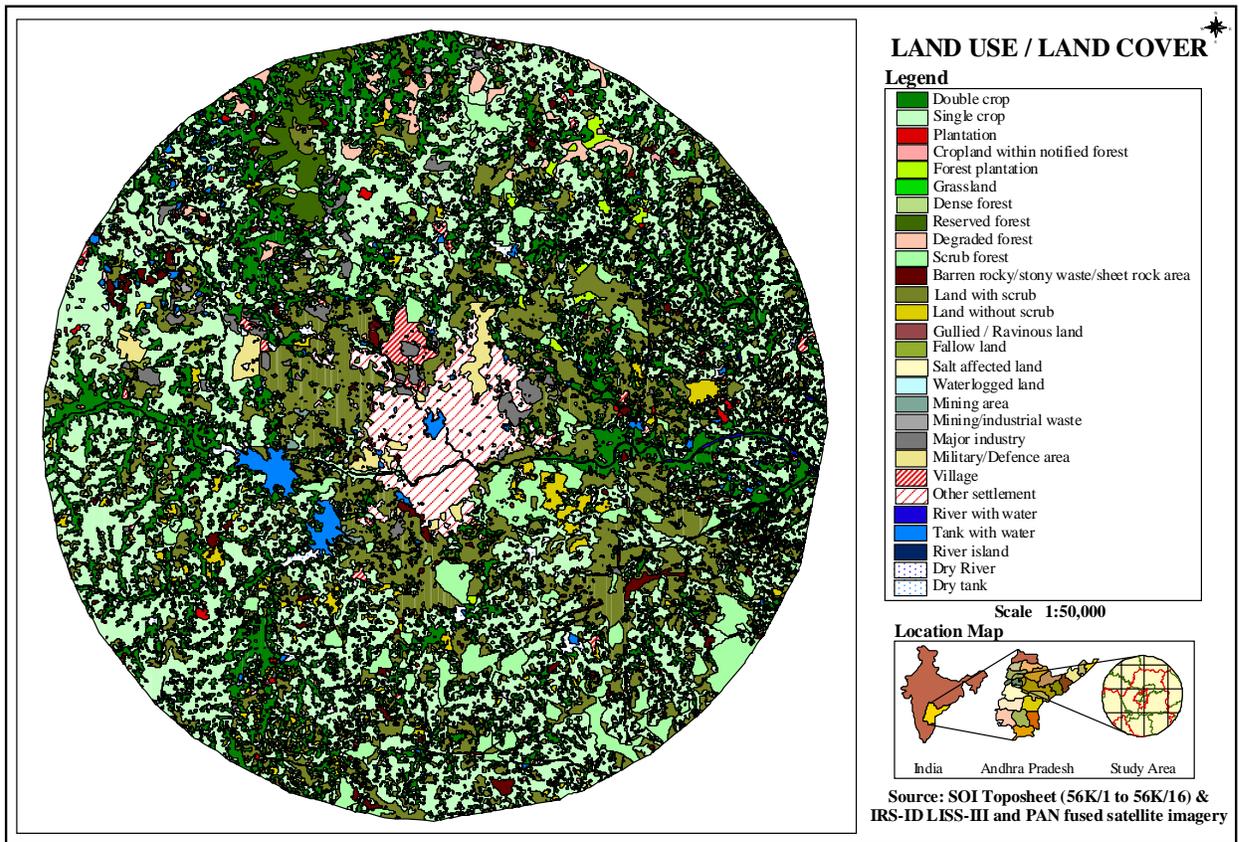


Figure 3: Land Use/ Land Cover map of the study area

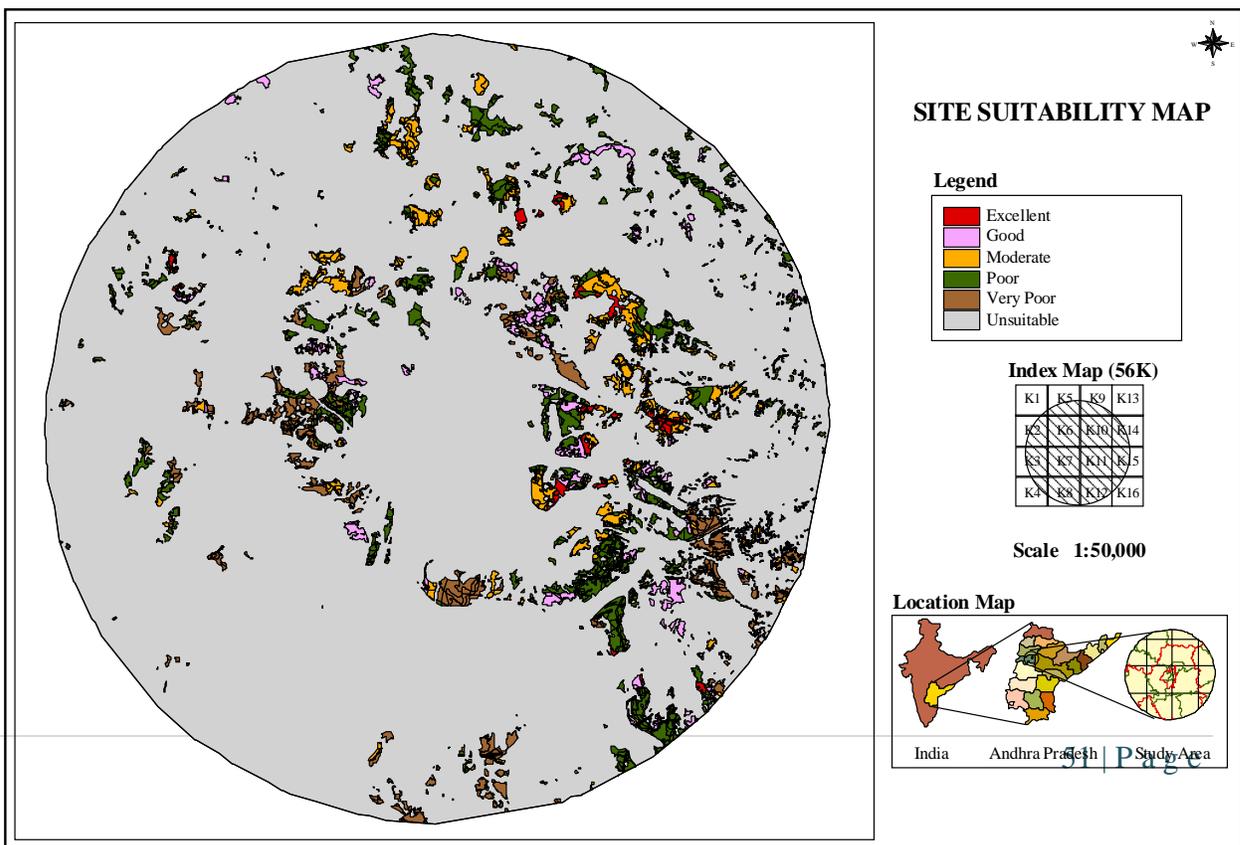


Figure 4: Site suitability map

An improved Particle Swarm Optimization

¹ V.Selvi,

² J.Saigeetha

³ Dr.R.Umarani

¹ Lecturer in CS Dept, Nehru Memorial College, Puthanampatti.Trichy(dt),TamilNadu.

² Lecturer in CS Dept, Nehru Memorial College, Puthanampatti.Trichy(dt),TamilNadu.

³Associate Professor in Computer Science, Sri Sarada College for Women, Salem, Tamil Nadu.

Abstract

Adaptive inertia weight is proposed to rationally balance the global exploration and local exploration abilities for particle swarm optimization. The result algorithm is called inertia particle swarm optimization (AIW-PSO). To avoid the premature convergence caused by basic Particle Swarm Optimization (PSO), a new particle Swarm Optimization algorithm with adaptive inertia weight (AIW –PSO) is proposed. Inertia weight is adaptively changed according to the algorithm. This paper describes a method for improving the final accuracy and the convergence speed of Particle Swarm Optimization (PSO) by adapting its inertia factor in the velocity updating equation and also by adding a new coefficient. In order to demonstrate the effectiveness of AIW-PSO, comprehensive experimental were conducted on three well-known benchmark functions.

Key word: PSO, Inertia weight, Benchmark functions

1.Introduction:

As product and engineering desing becomes more and more complicated, the objective function of optimization design is increasingly high dimensional, non-convex, and highly nonlinear. Traditional optimization design method usually operate difficultly and ineffectively, easily obtain local optimal. Therefore, finding a simple optimization method that can obtain global optimal quickly and effectively has an important significance to engineering design optimization.

In recent years, with the development and widely application of evolution algorithm and intelligent algorithm, particle swarm optimization (PSO) has been proven to be a better global optimization method with simple operation and parallel search.

In population-based optimization methods, proper control of global exploration and local exploitation is crucial in finding the optimum solution efficiently. Shi and Eberhart introduced the concept of inertia weight to the original version of PSO, in order to balance the local and global search during the optimization process.

1.1 Some significant variants of the classical PSO

Since its introduction by Kennedy and Eberhart in 1995, PSO has been subjected to empirical and theoretical investigations by several researchers. Shi and Eberhart introduced a new parameter ω , now well-known as *inertia weight*, to the original version of PSO in the following way:

$$V_{i,d}(t) = \omega \cdot V_{i,d}(t-1) + c_1 \cdot \text{rand}_1 \cdot (pbest_{i,d} - x_{i,d}(t-1)) + c_2 \cdot \text{rand}_2 \cdot (gbest_d - x_{i,d}(t-1)) \dots (1)$$

The inertia weight is used to balance the global and local search abilities. A large inertia weight is more appropriate for global search and a small inertia weight facilitates local search.

1.2 The Inertia-adaptive PSO Algorithm

Premature convergence occurs when the positions of the most of the particles of the swarm stop changing over successive iterations although the global optimum remains undiscovered. This may happen if the swarm uses a small inertia weight or a constriction coefficient. From the basic equations of PSO, we see that if $v_{i,d}$ is small and in addition to that $|pbest_{i,d} - x_{i,d}|$ and $|gbest_d - x_{i,d}|$ are small enough, $v_{i,d}$ cannot attain a Large value in the upcoming generations. That would mean a loss of exploration power. This can occur even an early stage of the search process, when the particle itself is the global best causing $|pbest_{i,d} - x_{i,d}|$ and $|gbest_d - x_{i,d}|$ to be zero and, gets damped quickly with the ratio ω . Also the swarm suffers from loss of diversity in later generations if $pbest$ and $gbest$ are close enough.

In this work we incorporate two modifications into the classical PSO scheme which prevent false convergence and helps provide excellent quality of final result without imposing any serious burden in terms of excess number of function evaluations (FEs). The first of these modifications involves modulation of the inertia factor ω according to distance of the particles of a particular generation from the global best. The value of ω for each particle is given by:

$$\omega = \omega_0 \cdot (1 - (dist_i / \max_dist)) \dots \dots \dots (2)$$

where ω_0 (0.5,1), $dist_i$ is the curredcnt Euclidean distance of i -th particle from the global best i . e.

D

$$\text{dist}_i = \left(\sum_{d=1}^n (\text{gbest}_d - x_{i,d})^2 \right)^{1/2} \dots\dots\dots(3)$$

and max_dist is the maximum distance of a particle from the global best in that generation i.e.

$$\text{max_dist} = \arg_i \max(\text{dist}_i) \dots\dots\dots(4)$$

This modulation of the inertia factor ensures that in case of particles that have moved away from the global best, the effect of attraction towards global best will predominate. To avoid premature convergence this we must ensure that the particle has mobility in the later stages. In order to achieve our purpose, the position update equation is modified as follows:

$$X_{i,d}(t) = (1 - \rho) \cdot X_{i,d}(t - 1) + V_{i,d}(t) \dots\dots\dots(5)$$

where ρ is a uniformly distributed random number in the range (-0.25, 0.25). From now on, we shall refer to this new algorithm as IAPSO (Inertia-adaptive PSO).

2. Experimental Results

Table1. Benchmark functions

| Name of the functions | Mathematical representation | Range of search | Range of initialization | Vmax |
|----------------------------|--|-----------------|-------------------------|------|
| Sphere function | $f_1(x) = \sum_{i=1}^n x_i^2$ | $(-100, 100)^n$ | $(50, 100)^n$ | 100 |
| Rosenbrock function | $f_2(x) = \sum_{i=1}^n [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$ | $(-100, 100)^n$ | $(15, 30)^n$ | 100 |
| Rastrigrin function | $f_3(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x) + 10]$ | $(-10, 10)^n$ | $(2.56, 5.12)^n$ | 10 |

population size=50 ,Iteration=500, dimension=10

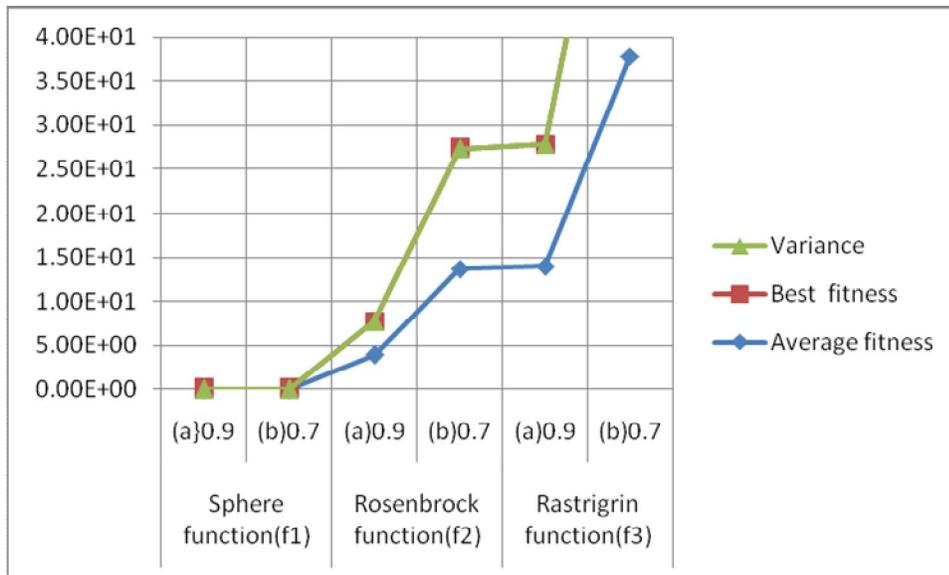
Table2. Various factors based on inertia weight

| Name of the function | Inertia weight | Average fitness | Best fitness | Variance |
|----------------------|----------------|-----------------|--------------|----------|
| | | | | |

| | | | | |
|-------------------------|--------|--------------|--------------|---------------|
| Sphere function(f1) | (a)0.9 | 2.0261519E-8 | 2.0261043E-8 | 3.6022766E-14 |
| | (b)0.7 | 5.2666E-7 | 5.2666E-7 | 2.8215297E-16 |
| Rosenbrock function(f2) | (a)0.9 | 3.8576868 | 3.8566718 | 2.562219E-5 |
| | (b)0.7 | 13.6800995 | 13.680098 | 7.944004E-7 |
| Rastrigrin function(f3) | (a)0.9 | 13.929418 | 13.929418 | 6.686148E-11 |
| | (b)0.7 | 37.808346 | 37.808346 | 3.9099223E-15 |

The inertia weight is mainly used to improve local and global search facilities. The best fitness value can be changed according to the inertia weight. The inertia weight has been reduced the best fitness value has been increased. To improve optimization, the population value can be adjusted based on inertia weight.

Graph 1:



3. Conclusion

This paper describes a method for improving speed and optimize the performance in local and global exploration for PSO. By this fair strategy, dynamically adjusted inertia weight, the performance of PSO algorithm could be improved. The experiments were conducted using different bench mark functions with various dimensions.

References:

- 1.J. Kennedy, “Stereotyping: Improving particle swarm performance with cluster analysis,” in *Proc. IEEE Int. Conf. Evolutionary Computation*, vol. 2, 2000, pp. 303–308.
2. . Lovbjerg and T. Krink, “Extending particle swarm optimizers with self-organized critically,” in *Proc. IEEE Int. Congr. Evolutionary Computation*, vol. 2, Honolulu, HI, May 2002, pp. 1588–1593.
- 3 .K. E. Parsopoulos and M. N. Vrahatis, “Particle swarm optimization method in multi objective problems,” in *Proc. ACM Symp. Applied Com-puting 2002 (SAC 2002)*, 2002, pp. 603–607.
- 4.M. S. Arumugam and M. V. C. Rao, “On the performance of the particle swarm optimization algorithm with various inertia weight variants for computing optimal control of a class of hybrid systems,” *Discrete Dynamics in Nature and Society*, Article ID 79295, 17 pages, 2006.
5. Y. del Valle, G. K. Venayagamoorthy, S. Mohagheghi, J.-C. Hernandez, and R. G. Harley, “Particle swarm optimization: basic concepts, variants and applications in power systems,” *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 2, pp. 171–195, 2008.
6. M. A. Abido, “Multiobjective particle swarm optimization technique for environmental/economic dispatch problem,” *Electric Power System Research*, vol. 79, no. 7, pp. 1105–1113, 2009.

NEURAL NETWORKS TOWARDS MEDICAL

Prof. A. Maithili¹

Dr. R. Vasantha Kumari²

Mr. S. Rajamanickam³

*(Director, Master of Computer Applications, Pope John Paul II College of Education, Puducherry, India)

** (Principal, KWC, Puducherry, India)

*** (Research Scholar, Pope John Paul II College of Education, Puducherry, India)

ABSTRACT

The Neural Networks are best at identifying patterns or trends in data and they are well suited for predicting or forecasting. Hence neural networks are extensively applied to biomedical systems. An analysis is carried out to motivate neural network applications in medical diagnosis. A special note is made on neural network effort on cancer diagnosis. This paper focuses on the importance of application of neural networks in the medical world, particularly in the diagnosis of various diseases. Technical concepts like MLP, SVM, RBF, CANFIS, TLRN, PCA, SOM, and SNNS are applied for diagnosing various diseases using neural networks.

Keywords - ANN, MLP, TLRN, RBF, CANFIS, SNNS.

I. INTRODUCTION

The term neural network was traditionally refers to a network of biological neurons. The modern usage of the term refers to artificial neural networks, which are composed of artificial neurons. Thus the term has two distinct usages, Biological neural networks and Artificial Neural Networks.

Biological neural networks are made up of real biological neurons that are connected in the peripheral nervous system or central nervous system. In the field of neuroscience, they are identified as a group of neurons that performs a specific physiological function in the laboratory analysis. Artificial neural networks are composed of interconnecting artificial neurons. This network may either be used to gain an understanding of biological neural networks, or for solving artificial intelligence problems.



Fig.1 Neural Networks

The neural networks exhibit mapping capability, they can map input pattern to their associated output patterns. They can learn by examples. Neural network architectures can be

trained with known examples of a problem before they are tested for their inference capacity on unknown instances of the problem. They can identify new objects previously untrained. They possess the capacity to generalize. Thus, they can predict new outcomes from past trends. They are robust systems and fault tolerant. They can recall full patterns from incomplete, partial or noisy patterns. They can also process information in parallel, at high speed and in a distributed manner.

II. ARCHITECTURE OF NEURAL NETWORKS

1. Feed-forward networks:

Feed-forward ANNs allow signals to travel one way only; from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. This type of organization is also referred to as bottom-up or top-down.

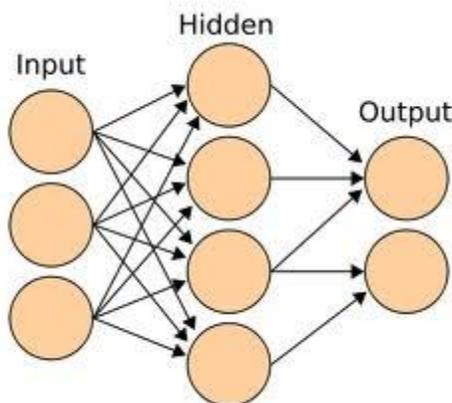


Fig.2 Feed-Forward Networks

2. Feedback networks:

Feedback networks can have signals travelling in both directions by introducing loops in the network. Feedback networks are very powerful and can get extremely complicated. Feedback networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations.

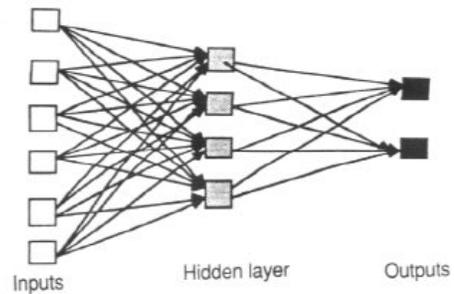


Fig.3 Feedback Networks

III. NEURAL NETWORKS LAYER

The common type of artificial neural network consists of three groups, or layers, of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. The activity of the input units represents the raw information that is fed into the network. The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units. The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units.

This simple type of network is interesting because the hidden units are free to construct their

own representations of the input. The weights between the input and hidden units determine when each hidden unit is active, and so by modifying these weights, a hidden unit can choose what it represents. We can also distinguish single-layer and multi-layer architectures. The single-layer organization, in which all units are connected to one another, constitutes the most general case and is of more potential computational power than hierarchically structured multi-layer organizations. In multi-layer networks, units are often numbered by layer, instead of following a global numbering.

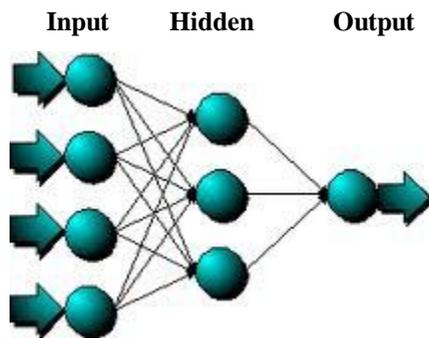


Fig.4 Neural Networks Layer

IV. APPLICATIONS OF NEURAL NETWORKS

Neural Networks have broad applications to the real world business problems. They have already been successfully applied in many industries. Since neural networks are best at identifying patterns or trends in data, they are well suited for prediction or forecasting. These include Sales forecasting, Industrial process control, Customer research, Data validation, Risk management, Target marketing.

V. NEURAL NETWORKS IN MEDICINE

The Artificial Neural Networks (ANN)

receives extensive application to biomedical systems in the next few decades. The research is focusing on modelling parts of the human body and recognising diseases from various scans (e.g. cardiograms, CAT scans, ultrasonic scans, etc).



Fig.5 Neural Networks in Medicine

Neural Networks are ideal in recognising diseases using scans there is no need to provide a specific algorithm on how to identify the diseases. Neural Networks learn by example so the details of how to identify the diseases are not needed. What is needed is a set of examples that are comprehensive of all the variations of the diseases. The examples need to be selected very carefully if the system is to perform reliably and efficiently.

1. Modelling and Diagnosing the Cardiovascular System:

Neural Networks are used experimentally to model the human cardiovascular system. Diagnosis can be achieved by building a model of the cardiovascular system of an individual and comparing it with the real time physiological measurements taken from the patient. If this routine is carried out regularly, potential harmful medical conditions can be detected at an early stage and thus

make the process of combating the disease much easier.



Fig.6 Cardio Vascular System

A model of an individual's cardiovascular system must mimic the relationship among physiological variables (i.e., heart rate, systolic and diastolic blood pressures, and breathing rate) at different physical activity levels. If a model is adapted to an individual, then it becomes a model of the physical condition of that individual. The simulator will have to be able to adapt to the features of any individual without the supervision of an expert. This calls for a neural network.

Another reason that justifies the use of ANN technology is the ability of ANNs to provide sensor fusion which is the combining of values from several different sensors. Sensor fusion enables the ANNs to learn complex relationships among the individual sensor values, which would otherwise be lost if the values were individually analyzed. In medical modelling and diagnosis, this implies that even though each sensor in a set may be sensitive only to a specific physiological variable, ANNs are capable of detecting complex medical conditions by fusing the data from the individual biomedical sensors.

2. Electronic Noses:

ANNs are used experimentally to implement electronic noses. Electronic noses have several potential applications in telemedicine. Telemedicine is the practice of medicine over long distances via a communication link. The electronic nose would identify odours in the remote surgical environment. These identified odours would then be electronically transmitted to another site where an odor generation system would recreate them. Because the sense of smell can be an important sense to the surgeon, telesmell would enhance telepresent surgery.



Fig.7 Electronic Nose

3. Instant Physician:

An application developed in the mid-1980s called the "instant physician" trained an auto associative memory neural network to store a large number of medical records, each of which includes information on symptoms, diagnosis, and treatment for a particular case. After training, the net can be presented with input consisting of a set of symptoms; it will then find

the full stored pattern that represents the "best" diagnosis and treatment.



Fig.8 Instant Physician

VI. ADVANCEMENT OF NEURAL NETWORKS IN MEDICAL SCIENCE

1. Examine common characteristics in large volume of data:

Locating common characteristics in large amounts of data is a type of classification problem. Neural networks can be used to solve classification problems, typically through Multi-Layer Perceptron (MLP) and Support Vector Machines (SVM) type networks.

Examples of classification applications in medicine include dividing research populations or data into groups for further study. For example, data from studies of body movement could be classified into different patterns to aid with physical therapy.

2. Predicting results based on existing data:

Forecasting results based on existing data is a type of function approximation problem. Neural networks can be used to solve function approximation problems, typically through Multi-Layer Perceptron (MLP), Radial Basis Function

(RBF) and CANFIS (Co-Active Neuro-Fuzzy Inference System) type networks.

Examples of function approximation applications in medicine include the prediction of patient recovery and automated changes to device settings. For example, data from studies of potential recovery level of patients can provide realistic estimates to patients while helping facilities cut costs by better allocating resources.

3. Predict the progression of medical data over time:

Predicting the progression of medical data over time is a type of time-series prediction problem. Neural networks can be used to solve time-series problems, typically through Time-Lagged Recurrent (TLRN) type network.

Examples of time-series predictions in medicine include the prediction of cell growth and disease dispersion. For example, data from studies of muscle stimulation patterns of arm movements can be used to control mouse movements on a computer screen.

4. Identify specific characteristics in medical imagery:

Identifying specific characteristics in medical imagery is a type of image processing problem. Neural networks can be used to solve image processing problems, typically through Principal Component Analysis (PCA) type network.

Examples of image processing in medicine include the detection of characteristics in ultrasound and x-ray features. For example, image data from studies of mammograms can be used for the detection of breast cancer.

5. Group medical data based on key characteristics:

Grouping of medical data based on key characteristics is a type of clustering problem. Neural networks can be used to solve clustering problems, typically through Self-Organizing Map (SOM) type network.

Examples of clustering in medicine include the detection of key characteristics in demographics or pre-existing conditions. For example, data from studies combined with sensitivity analysis can reverse engineer a biologically plausible relationship from real world data.

VII. CANCER

Cancer is primarily an environmental disease, though genetics influence the risk of some cancers. Common environmental factors leading to cancer include: tobacco use, poor diet and obesity, infection, radiation, lack of physical activity, and environmental pollutants. These environmental factors cause or enhance abnormalities in the genetic material of cells. Cell reproduction is an extremely complex process that is normally tightly regulated by several classes of genes, including oncogenes and tumor suppressor genes. Hereditary or acquired abnormalities in these regulatory genes can lead to the development of cancer. A small percentage of cancers, approximately five to ten percent, are entirely hereditary.

The presence of cancer can be suspected on the basis of clinical signs and symptoms, or findings after medical imaging. Definitive diagnosis of cancer, however, requires the microscopic examination of a biopsy specimen. Most cancers can be treated, with the most important modalities being chemotherapy, radiotherapy and surgery. The prognosis in cancer cases can be greatly influenced by the type and location of the cancer and the extent of disease.

While cancer can affect people of all ages, and a few types of cancer are more common in children than in adults, the overall risk of developing cancer generally increases with age, at least up to age 80-85 yr. In 2007, cancer caused about 13% of all human deaths worldwide (7.9 million). Rates are rising as more people live to an old age and as mass lifestyles changes occur in the developing world.

1. Origin of Cancer:

The organs in our body are made up of cells. Cells divide and multiply as the body needs them. When these cells continue multiplying when the body doesn't need them, the result is a mass or growth, also called a tumor. These growths are considered either benign or malignant. Benign is considered non-cancerous and malignant is cancerous. Benign tumors rarely are life threatening and do not spread to other parts of the body. They can often be removed. Malignant tumors, however, often invade nearby tissue and organs, spreading the disease.

2. Spreading of Cancer:

The cells within malignant tumors have the ability to invade neighboring tissues and organs, thus spreading the disease. It is also possible for cancerous cells to break free from the tumor site and enter the bloodstream, spreading the disease to other organs. This process of spreading is called metastasis. When cancer has metastasized and has affected other areas of the body, the disease is still referred to the organ of origination. For example, if cervical cancer spreads to the lungs, it is still called cervical cancer, not lung cancer. Although most cancers develop and spread this way -- via an organ - blood cancer like leukemia do not. They affect the blood and the

organs that form blood and then invade nearby tissues.

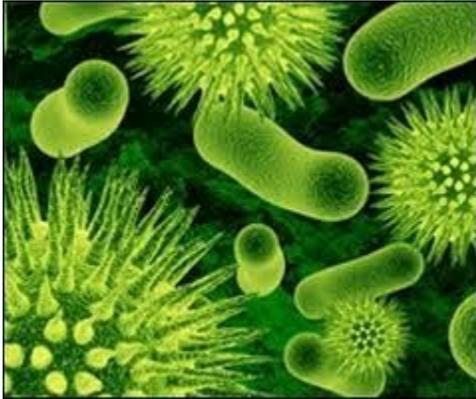


Fig.9 Spreading of Cancer Cells

VIII. DIAGNOSING CANCER USING NEURAL NETWORKS

There are several systems available for the diagnosis and selection of therapeutic strategies in breast cancer. A neural network judged the possible recurrence rate of tumors correctly in 960 of 1008 cases by using data from lymphatic node positive patients (tumor size, number of palpable lymphatic nodules, tumor hormone receptor status, etc.). Baker et al. reported that they came to similar results by neural network evaluation of the parameters of the BI-RADS standardized code system. Fogel stated in his paper on neural network recognition of breast cancer that evaluation of mammographic, cytological and epidemiological findings in an integrated system are thought to be useful in the diagnostic process.

IX. CONCLUSION

In this article efforts are taken to analyse the application of Artificial Neural Network towards medical diagnosis. Step by step analysis is carried out

and emphasis is made on diagnosis of cancer. Identifying diseases helps Doctors to save human to a major extent. Mathematical modelling of neural networks in analyzing their medical diagnosis using allowable range of parameters is taken as future work.

FURTHER ENHANCEMENT

The field of Bio-Medical Engineering is an emerging branch where attentions are called for every individual. The present world which is adversely affected by Tsunami and Earthquakes faces the terrific attack of diseases and it is of immense need for protecting against as well as getting cured from threatening diseases.

Care is taken to apply neural networks in diagnosing recent diseases like Swine Flu, Chicken Guinea, Brain tumor, and Brain fever and so on.

Mathematical modeling of neural networks will be focused on and various parameters will be engaged so as to get the required result to desired degree of accuracy.

REFERENCES

Journal Papers:

[1]Neural Networks at Pacific Northwest National Laboratory

<http://www.emsl.pnl.gov:2080/docs/cie/neural/neural.homepage.html>

[2]Artificial Neural Networks in Medicine

<http://www.emsl.pnl.gov:2080/docs/cie/techbrief/NN.techbrief.html>

[3] Electronic Noses for Telemedicine
<http://www.emsl.pnl.gov:2080/docs/cie/neural/papers/2/keller.ccc95.abs.html>

[4] Pattern Recognition of Pathology Images
<http://kopernik-eth.npac.syr.edu:1200/Task4/pattern.html>

[5] Ageing or cancer: a review on the role of caretakers and gatekeepers.

D Van Heemst, P M Den Reijer, R G J Westendorp in European journal of cancer Oxford England 1990 (2007)

[6] Tumor suppressor genes in breast cancer: the gatekeepers and the caretakers.

Andre M Oliveira, Jeffrey S Ross, Jonathan A Fletcher in American Journal of Clinical Pathology (2005)

Books:

[7] An introduction to neural computing. Aleksander, I. and Morton, H. 2nd edition

[8] Neural Networks by Eric Davalo and Patrick Naim 2nd edition

[9] Learning internal representations by error propagation by Rumelhart, Hinton and Williams (1986). 1st edition

[10] Recessive human cancer susceptibility genes M Wada, J Yokota in Nippon Rinsho (1988) 1st edition

Theses:

[11] Industrial Applications of Neural Networks (research reports Esprit, I.F.Croall, J.P.Mason)

[12] Klimasauskas, CC. (1989). The 1989 Neuro Computing Bibliography. Hammerstrom, D. (1986). A Connectionist/Neural Network Bibliography.

[13] DARPA Neural Network Study (October, 1987-February, 1989). MIT Lincoln Lab. Neural Networks, Eric Davalo and Patrick Naim

[14] Assimov, I (1984, 1950), Robot, Ballantine, New York.

[15] Ageing or cancer: a review on the role of caretakers and gatekeepers.

D Van Heemst, P M Den Reijer, R G J Westendorp in European journal of cancer Oxford England 1990 (2007)

Studies on production of lactic acid from various wastes using *Lactobacillus rhamnosus* and *Lactococcus lactis subsp lactis*

S. Bhuvaneshwari*, and V. Sivasubramanian

Department of Chemical engineering
National Institute of Technology Calicut, Kerala 673 601, India

Abstract

Lactic acid can be produced by fermentation of organic matters present in various wastes with the addition of various carbon and nitrogen sources using *Lactobacillus* species. Organic wastes from kitchen garbage and food-processing industries are especially rich in carbon and moisture content. The present study investigates the treatment of the organic wastes using microbiological process for effective usage of waste and to develop value added products from it. The organic wastes used in this processes are domestic wastes, vegetable wastes, fruit wastes, bakery wastes and whey. The microorganisms used in the synthesis of lactic acid are *Lactobacillus rhamnosus* and *Lactococcus lactis subsp lactis*. The fermentation study includes the effect of various wastes, effect of carbon and nitrogen sources with optimal waste medium, various concentrations of optimal carbon and nitrogen source and effect of time for free cells and their immobilized form on lactic acid production and biomass growth for both the organisms. The optimal production of lactic acid was determined as 35.45 g/l and bacterial growth was 1.34 g/l from whey by *Lactobacillus rhamnosus*. Optimal carbon source was mannitol, which yields 38.3 g/l of lactic acid and 1g/l of microbial growth. The study on effect of various concentrations of mannitol resulted a higher yield of lactic acid production at a concentration of 12 g/l with the production of 40.85 g/l. whereas the biomass yield was about 1.12 g/l at this optimum concentration. 12g/l of peptone as nitrogen source effectively yields 53.5g/l of lactic acid and cell growth of 1.56 g/l. The results are promising for higher yield of lactic

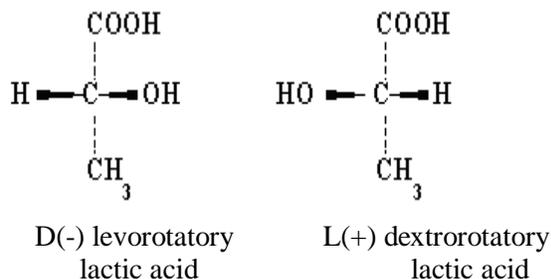
acid to be produced using *Lactobacillus* species compared to *Lactococcus lactis subsp lactis*.

Key words: Carbon sources, Lactic acid, *Lactococcus lactis subsp lactis*, *Lactobacillus rhamnosus*, and Nitrogen sources.

1. INTRODUCTION

Lactic acid is one of the most widely used organic acids in food industry. Recently, lactic acid has gained much interest because it can be used as a raw material for the production of biodegradable polymers with applications in medical, pharmaceutical, and food-packaging industries. Lactic acid, is a valuable industrial chemical used as an acidulant, preservative in the food industry, pharmaceutical, leather and textile industries as well as a chemical feedstock. Lactic acid bacteria have the property of producing lactic acid. The genera *Bacillus*, *Leuconostoc*, *Pediococcus* and *Streptococcus* are important members of this group. The species which have been used for lactic acid production are *L. sporogenes*, *L. acidophilus*, *L. plantarum*, *L. casei*, *L. brevis*, *L. delbrueckii*, *L. lactis*. During these days, food-processing industry generates large volumes of carbohydrate wastes, which are ideal substrates for lactic acid production. Lactic acid is then converted into useful, high-value products such as lactic acid derivatives, which include biodegradable plastics, oxy chemicals, "green" solvents, and specialty chemicals[1]. The general equation for lactic acid fermentation is $C_6H_{12}O_6 \rightarrow 2 C_3H_6O_3$. The levels of optical isomeric forms of lactic acid produced depend upon the nature of the culture [2]. The

structural configurations of these isomers are as follows :



Generally the optimum pH and temperature for a batch culture of *Lactobacillus sp.* was found to be 6.0 and 36°C, respectively. For *L. acidophilus* pH was 6.5 or 8.0 and temperature was 37 °C, *Lactobacillus pentosus* the pH was 7.6 and temperature was 32 °C. *Lactobacillus delbrueckii* utilized industrial corn cob waste from xylose manufacturing and the yield of lactic acid from glucose were 48.7 g l⁻¹ and 95.2% respectively. *Lactobacillus delbrueckii* NRRL B445 yielded lactic acid from pulp mill solid waste ranged from 86 to 97% [3]. Efficient production of lactic acid was obtained from raw starch by *Streptococcus bovis* 148[4]. *Lactobacillus sake* & *Lactobacillus casei* yielded the lactic acid conversion of 48% and 56% respectively however lactic acid conversion increased to 71% by co- inoculation of both the strains from soya bean stalk. Lactic acid production was investigated using filamentous fungus *Rhizopus arrhizus* and the yield was only 75% and 61% results from the glucose medium. *Rhizopus arrhizus* produced lactic acid from waste potato starch[5]. All bacteria used for lactic acid production was more efficient in MRS broth[6].

Lactic acid bacteria (LAB) which are used throughout the world for manufacture of a wide variety of traditional fermented foods [7], are beneficial probiotic organisms that contribute to improved nutrition, microbial balance, and immuno-enhancement of the intestinal tract, as well as lower cholesterol along with that it will produce antimicrobial substances, sweeteners, sugar polymers, useful enzymes and aromatic compounds .Lactic acid bacteria foods and supplements have been found to modulate

inflammatory and hypersensitivity responses, an observation thought to be at least in part due to the regulation of cytokine function. Clinical studies suggest that they can prevent reoccurrences of Inflammatory Bowel Disease in adults, as well as improve milk allergies and decrease the risk of atopic eczema in children. *L. rhamnosus* appears to protect the urogenital tract by excreting biosurfactants to inhibit the adhesion of vaginal and urinary pathogens. *Lactobacillus* also releases acids, bacteriocins, and hydrogen peroxide which inhibit the pathogen growth. *Lactobacillus Rhamnosus* has also been shown beneficial in the prevention of rotavirus diarrhea in children. The use of *Lactobacillus rhamnosus* for probiotic therapy does present a risk of sepsis. *L. rhamnosus* adheres to the mucous membrane of the intestine and may help to restore the balance of the GI micro flora; promote gut-barrier functions; diminish the production of carcinogenic compounds by other intestinal bacteria; and activate the innate immune response and enhance adaptive immunity, especially during infections [8].

2. MATERIALS AND METHODS

2.1 Material:

All the chemicals used are of analytical grade and samples were prepared using deionized water. *Lactobacillus rhamnosus* MTCC1408, and *Lactococcus lactis subsp lactis* MTCC440 were purchased from MTCC (Microbial Type Culture collection, Chandigarh, India).

2.2 Methods:

Lactococcus lactis subsp lactis MTCC440 were cultured on medium containing skim milk (10%), tomato juice (10%) and yeast extract agar slants at pH of 7 for 48 hours at 37°C under aerobic conditions. Subculturing is done once in 30 days.

Lactobacillus rhamnosus MTCC1408 were cultured on MRS (Man, Rogosa and Sharpe medium) slants adjusted to pH 6.2-6.6 and incubated at 37°C for 48 hours under aerobic conditions. Sub culturing was done once in 30 days.

2.2.1 Media composition**Composition of MRS medium (Adjusted pH to 6.2-6.6)**

| | |
|---------------------------------------|----------|
| Peptone | - 10.0 g |
| Beef extract | - 10.0 g |
| Yeast extracts | - 5.0 g |
| Glucose | - 20.0 g |
| Tween 80 | - 1.0 ml |
| Na ₂ HPO ₄ | - 2.0 g |
| Sodium acetate | - 5.0 g |
| Triammonium citrate | - 2.0 g |
| MgSO ₄ , 7H ₂ O | - 0.2 g |
| MnSO ₄ , 4H ₂ O | - 0.2 g |
| Agar | - 15.0 g |
| Distilled water | - 1.0 l |

Composition of Potato Dextrose Broth medium (Adjusted pH to 5.6)

| | |
|-----------------------------|-----------|
| Potato (scrubbed and diced) | - 200.0 g |
| Dextrose | - 20.0 g |
| Distilled water | - 1.0 l |
| Agar | - 15 g |

2.2.2 Fermentation in various wastes

The various wastes such as domestic wastes, bakery wastes, fruit wastes, vegetable wastes and whey were used as fermentation medium. The domestic wastes containing kitchen garbage waste, agricultural wastes, gardening waste and stationary waste were collected from house hold and the organic matter were selected from the waste. These organic matters of about 1 kg was ground along with 1100 ml of distilled water and filtered to get the filtrate which is used as the fermentation medium.

Similarly vegetable waste, fruit waste and bakery waste were collected from household. One kg of each waste was ground along with 1100 ml of distilled water and filtered to obtain the filtrate which is used as the fermentation medium. All the above wastes will have different chemical compositions of carbohydrates, lignin, cellulose, hemicelluloses, protein and lipid. Whey was obtained from the household, since it was the by-product while preparing sweets and paneer.

All these five waste extracts were used as the fermentation medium. The two bacterial was inoculated in respective medium and incubated at pH of about 6.5 for 24 hours at 37°C. In 1000 ml of the sterilized wastes medium, 1% (10 ml) of 24 hours bacterial inoculum was added and incubated at pH of about 6.5 for 24 hours at 37°C. After 8 hour, the pH was maintained between 5.5-6 by addition of 2% CaCO₃. After fermentation the lactic acid concentration, biomass growth and substrate conversion was found for each organism in different wastes.

2.2.3 Study of various carbon sources

The optimal waste medium for the production of lactic acid was analysed for both organisms. Galactose, glucose, lactose, maltose and mannitol were selected as the carbon sources. The lactic acid concentration, biomass growth and substrate conversion for each organism were found in the best suited waste medium containing 10g/l of each carbon sources.

2.2.4 Study of various concentrations of the optimal carbon source

The carbon source which gives maximum yield of lactic acid for each organism was estimated. The various concentrations of the carbon sources such as 6, 8, 10, 12, 14 g/l are added to the optimal waste medium and lactic acid productivity is analyzed.

2.2.5 Study of various nitrogen sources

The optimal waste medium for the production of lactic acid is analyzed for each organisms. Sodium nitrate, peptone, potassium nitrate, urea and yeast extract were selected as the nitrogen sources. The lactic acid concentration, biomass growth and substrate conversion for each organism were found in the best suited waste medium containing 10g/l of each nitrogen sources.

2.2.6 Study of various concentrations of the optimal nitrogen source

The various nitrogen sources which produces maximum yield of lactic acid for each microorganisms was estimated. The various

concentrations of the nitrogen sources such as 6,8,10,12,14 g/l are added to the optimal waste medium and lactic acid productivity is analyzed.

2.2.7 Time profile of the microorganisms

The bacterial strains were grown on MRS agar plates by the method of four quadrants streaking using a sterile inoculum loop. For time profile study a single isolated colony was picked up from its plate from the help of a sterile inoculation loop and added to 10 ml of sterile MRS broth in a test tube to grow overnight (18h) in order to activate the colony taken from the plate. From the suspension of overnight grown bacterial culture, 10ml (1% inoculum) suspension was transferred in 1000 ml of sterile waste medium containing the best suited carbon or nitrogen sources in its optimal concentration.

3. RESULTS AND DISCUSSION

3.1 *Lactobacillus rhamnosus* MTCC1408 was inoculated in various wastes and MRS medium and incubated. Among the wastes as substrate, lactic acid production was found to be high in whey about 35.45 g/l and in the control MRS medium about 55.6 g/l. The biomass concentration was found to be high in whey about 1.34 g/l compared to other wastes and 1.59 g/l in MRS medium. Since the lactic acid production was comparatively high in whey than other wastes as illustrated in Table 1, further studies on effect of various carbon sources and nitrogen sources in lactic acid production was carried out with whey as substrate. Among the various carbon sources such as galactose, glucose, lactose, maltose and mannitol, *Lactobacillus rhamnosus* MTCC1408 showed higher yield of lactic acid about 38.3 g/l in mannitol. The biomass growth was found to be high in mannitol about 1 g/l as illustrated in Table 2. So the effect of various concentrations of mannitol were analysed and the higher yield of lactic acid production was found in 12 g/l concentration with the production of 40.85 g/l and the biomass yield was 1.12 g/l as illustrated in Table 3. The effect of various nitrogen sources such as sodium nitrate, peptone, potassium nitrate, urea and yeast extract on the

production of lactic acid by *Lactobacillus rhamnosus* MTCC1408 were studied and found to be high in peptone about 51.3 g/l of lactic acid. The biomass yield was found to be high in peptone of 1.43 g/l as illustrated in the Table 4. So the lactic acid production on various concentrations of peptone was studied and found maximum yield of lactic acid was in the concentration of 12 g/l which resulted in a maximum yield of 53.5 g/l of lactic acid. The biomass growth was found to be high in 12 g/l of peptone and the biomass produced was about 1.56 g/l as illustrated in the Table 5. On examining the time profile of *Lactobacillus rhamnosus* MTCC1408 the maximum lactic acid production and biomass growth was found at 48 hours and further increase in time showed no significant effect on lactic acid production and biomass yield, as shown in the Fig. 1 and Fig. 2.

3.2 *Lactococcus lactis subsp lactis* MTCC440

was inoculated in various wastes and with MRS medium and then incubated. Among the various wastes, lactic acid production was found to be high in whey of about 8.28 g/l and in the control MRS medium of about 42.2 g/l. The biomass concentration was found to be high in whey of about 0.74 g/l compared to other waste and 1.34 g/l in MRS medium, as illustrated in the Table 6. Since the lactic acid production was comparatively high in whey than other wastes, further studies on effect of various carbon sources and nitrogen sources was studied with whey as substrate. Among the various carbon sources such as galactose, glucose, lactose, maltose and mannitol, *Lactococcus lactis subsp lactis* MTCC440 showed higher yield of lactic acid about 13.2 g/l in glucose. The biomass growth was found to be high in glucose of about 0.97 g/l as illustrated in the Table 7. So the effect of various concentrations of glucose was further studied and found higher yield of lactic acid production was in 6 g/l concentration with the production of 24.2 g/l. The biomass yield was found to be high in 6 g/l of about 1.15 g/l as illustrated in the Table 8.

The effect of various nitrogen sources such as sodium nitrate, peptone, potassium nitrate, urea and yeast extract on the production of lactic acid

by *Lactococcus lactis subsp lactis* MTCC440 was studied and found to be high in sodium nitrate about 9.4 g/l of lactic acid and the biomass growth was found to be high in sodium nitrate about 0.95 g/l as illustrated in the Table 9. So the lactic acid production on various concentrations of sodium nitrate was studied and maximum yield of lactic acid was found in the concentration of 10 g/l, which resulted in a maximum yield of 9.4 g/l of lactic acid. The biomass growth was high in 10 g/l of sodium nitrate and the biomass produced was about 0.95 g/l as illustrated in the Table 10. On examining the time profile of *Lactococcus lactis subsp lactis* MTCC440 the maximum lactic acid production and biomass growth was found at 48 hours as shown in the Fig. 3 and Fig. 4.

3.3 Immobilized *Lactobacillus rhamnosus* MTCC1408 and *Lactococcus lactis subsp lactis* MTCC440

The immobilized cells were allowed for 5 runs of batch fermentation. During the first batch of fermentation, the lactic acid production was high and that increased further in second & third run of batch fermentation for both *Lactobacillus rhamnosus* and *Lactococcus lactis subsp lactis*. But the production was significantly decreased during fourth and fifth run of batch fermentation for both the species as illustrated in Fig. 5.

4. CONCLUSIONS

Based on the finding of the present study, it is concluded that the lactic acid production using *Lactobacillus rhamnosus* and *Lactococcus lactis subsp lactis* was high when whey was used as substrate. An additional supplement of glucose at a concentration of 6 g/l and sodium nitrate in a concentration of 10 g/l improved the production yield of lactic acid in case of *Lactococcus lactis subsp lactis*. Whereas *Lactobacillus rhamnosus* showed higher yield of lactic acid with mannitol as carbon source and with peptone as nitrogen source at an optimum concentration of 12 g/l. About 48 hours of fermentation time gave good yield of lactic acid for both the strains and on immobilization *Lactobacillus rhamnosus* found to have better

productivity of lactic acid compared to *Lactococcus lactis subsp lactis*. Hence, these strains could be good candidates for bioconversion of renewable sources for lactic acid production.

Nomenclature

| | |
|-----------|--------------------------------------|
| S_0 | - Initial feed concentration, g/l |
| S | - Final substrate concentration, g/l |
| X | - Biomass concentration, g/l |
| P | - Product concentration, g/l |
| $Y_{X/S}$ | - Substrate yield coefficient, g/g |
| $Y_{P/S}$ | - Product yield coefficient, g/l |
| SC | - Substrate conversion, % |
| DM | - Domestic waste |
| BW | - Baker waste |
| FM | - Fruit Waste |
| VM | - Vegetable waste |

Table 1

Lactic acid production by *Lactobacillus rhamnosus* MTCC1408 from various wastes and MRS medium

| Kinetic parameters | D W | B W | FW | VW | Whey | MRS |
|--------------------|------|------|------|------|------|------|
| S_0 (g/l) | 18.9 | 17.5 | 18.1 | 18.3 | 17.3 | 29.8 |
| S (g/l) | 15 | 13.6 | 14.1 | 13.9 | 12.7 | 14.1 |
| $S_0 - S$ (g/l) | 3.9 | 3.89 | 4.02 | 4.38 | 4.59 | 15.7 |
| SC (%) | 26 | 22.2 | 22.2 | 23.9 | 26.6 | 52.7 |
| X (g/l) | 0.8 | 1.11 | 1.18 | 1.21 | 1.34 | 1.59 |
| P (g/l) | 16 | 12.5 | 16.9 | 24.4 | 35.5 | 55.6 |
| $Y_{X/S}$ (g/g) | 0.21 | 0.28 | 0.29 | 0.28 | 0.29 | 0.10 |
| $Y_{P/S}$ (g/l) | 4.16 | 3.22 | 4.20 | 5.58 | 7.72 | 3.54 |

Table 2**Effect of various carbon sources on lactic acid production and biomass yield**

| Kinetic parameters | Galactose (g/l) | Glucose (g/l) | Lactose (g/l) | Maltose (g/l) | Mannitol (g/l) |
|--------------------------|-----------------|---------------|---------------|---------------|----------------|
| S ₀ (g/l) | 18.52 | 17.82 | 18.30 | 17.41 | 18.75 |
| S (g/l) | 14.41 | 13.67 | 13.77 | 12.92 | 14.12 |
| S ₀ - S (g/l) | 4.11 | 4.15 | 4.53 | 4.49 | 4.63 |
| SC (%) | 22.19 | 23.2 | 24.75 | 25.78 | 24.69 |
| X (g/l) | 0.86 | 0.89 | 0.95 | 0.93 | 1.0 |
| P (g/l) | 23 | 26.4 | 36.5 | 36.2 | 38.2 |
| Y _{x/s} (g/g) | 0.209 | 0.214 | 0.2097 | 0.207 | 0.215 |
| Y _{p/s} (g/l) | 5.596 | 6.361 | 8.057 | 8.062 | 8.25 |

Table 3**Effect of various concentrations of mannitol**

| Kinetic parameters | 6 (g/l) | 8 (g/l) | 10 (g/l) | 12 (g/l) | 14 (g/l) |
|--------------------------|---------|---------|----------|----------|----------|
| S ₀ (g/l) | 18.51 | 18.63 | 18.75 | 18.88 | 18.97 |
| S (g/l) | 14.12 | 14.08 | 14.12 | 14.1 | 14.56 |
| S ₀ - S (g/l) | 4.39 | 4.55 | 4.63 | 4.78 | 4.35 |
| SC (%) | 23.7 | 24.42 | 24.69 | 25.31 | 22.93 |
| X (g/l) | 0.94 | 0.97 | 1 | 1.12 | 0.89 |
| P (g/l) | 37 | 37.5 | 38.2 | 40.85 | 35 |
| Y _{x/s} (g/g) | 0.21 | 0.213 | 0.215 | 0.234 | 0.205 |
| Y _{p/s} (g/l) | 8.43 | 8.241 | 8.25 | 8.546 | 8.045 |

Table 4**Effect of various nitrogen sources on lactic acid production and biomass yield**

| Kinetic parameters | Sodium nitrate (g/l) | Peptone (g/l) | Potassium nitrate (g/l) | Urea (g/l) | Yeast extract (g/l) |
|--------------------------|----------------------|---------------|-------------------------|------------|---------------------|
| S ₀ (g/l) | 18.4 | 17.8 | 18.9 | 18.3 | 17.8 |
| S (g/l) | 14.1 | 12.3 | 14.57 | 14.2 | 13.1 |
| S ₀ - S (g/l) | 4.34 | 5.54 | 4.29 | 4.05 | 4.75 |
| SC (%) | 23.6 | 31.1 | 23.1 | 22.2 | 26.7 |
| X (g/l) | 1.03 | 1.43 | 0.97 | 0.89 | 1.11 |
| P (g/l) | 25 | 51.8 | 21.8 | 19 | 13.4 |
| Y _{x/s} (g/g) | 0.24 | 0.26 | 0.23 | 0.22 | 0.23 |
| Y _{p/s} (g/l) | 5.76 | 9.35 | 5.08 | 4.69 | 8.29 |

Table 5**Effect of various concentrations of peptone**

| Kinetic parameters | 6 (g/l) | 8 (g/l) | 10 (g/l) | 12 (g/l) | 14 (g/l) |
|--------------------------|---------|---------|----------|----------|----------|
| S ₀ (g/l) | 17.53 | 17.68 | 17.82 | 17.97 | 18.12 |
| S (g/l) | 12.42 | 12.48 | 12.28 | 12.28 | 12.81 |
| S ₀ - S (g/l) | 5.11 | 5.2 | 5.54 | 5.69 | 5.31 |
| SC (%) | 29.15 | 29.41 | 31.07 | 31.66 | 29.30 |
| X (g/l) | 0.87 | 1.01 | 1.43 | 1.56 | 1.31 |
| P (g/l) | 31.5 | 38.7 | 51.8 | 53.5 | 45.2 |
| Y _{x/s} (g/g) | 0.17 | 0.194 | 0.258 | 0.274 | 0.246 |
| Y _{p/s} (g/l) | 6.164 | 7.44 | 9.35 | 9.40 | 8.512 |

Table 6

Lactic acid production by *Lactococcus lactis* subsp *lactis* MTCC440 from various wastes and MRS medium

| Kinetic parameters | D W | B W | FW | VW | Whey | MRS |
|--------------------------|------|------|------|------|------|------|
| S ₀ (g/l) | 18.9 | 17.8 | 18.1 | 18.3 | 17.3 | 29.8 |
| S (g/l) | 15.2 | 14.1 | 14.1 | 14.6 | 13.2 | 16.1 |
| S ₀ - S (g/l) | 3.70 | 3.68 | 3.98 | 3.75 | 4.03 | 13.7 |
| SC (%) | 19.6 | 20.6 | 22.0 | 20.5 | 23.4 | 45.9 |
| X (g/l) | 0.65 | 0.63 | 0.73 | 0.67 | 0.74 | 1.34 |
| P (g/l) | 3.3 | 2.35 | 7.52 | 4.26 | 8.28 | 42.2 |
| Y _{x/s} (g/g) | 0.18 | 0.17 | 0.18 | 0.18 | 0.19 | 0.10 |
| Y _{p/s} (g/l) | 0.89 | 0.64 | 1.89 | 1.14 | 2.05 | 3.23 |

Table 7

Effect of various carbon sources lactic acid production and biomass yield

| Kinetic parameters | Galactose (g/l) | Glucose (g/l) | Lactose (g/l) | Maltose (g/l) | Mannitol (g/l) |
|--------------------------|-----------------|---------------|---------------|---------------|----------------|
| S ₀ (g/l) | 18.52 | 17.82 | 18.30 | 17.41 | 18.8 |
| S (g/l) | 14.38 | 13.24 | 14.09 | 13.22 | 14.7 |
| S ₀ - S (g/l) | 4.14 | 4.58 | 4.21 | 4.19 | 4.02 |
| SC (%) | 22.35 | 25.70 | 23 | 24.06 | 21.4 |
| X (g/l) | 0.78 | 0.97 | 0.87 | 0.83 | 0.71 |
| P (g/l) | 9.2 | 13.2 | 9.6 | 9.6 | 8.2 |
| Y _{x/s} (g/g) | 0.188 | 0.212 | 0.206 | 0.198 | 0.18 |
| Y _{p/s} (g/l) | 2.22 | 2.88 | 2.28 | 2.29 | 2.03 |

Table 8

Effect of various concentrations of glucose

| Kinetic parameters | 6 (g/l) | 8 (g/l) | 10 (g/l) | 12 (g/l) | 14 (g/l) |
|--------------------------|---------|---------|----------|----------|----------|
| S ₀ (g/l) | 17.63 | 17.71 | 17.82 | 17.95 | 18.07 |
| S (g/l) | 12.44 | 12.88 | 13.24 | 13.61 | 3.8 |
| S ₀ - S (g/l) | 5.19 | 4.83 | 4.58 | 4.34 | 4.27 |
| SC (%) | 29.43 | 27.27 | 25.70 | 24.17 | 23.63 |
| X (g/l) | 1.15 | 1.05 | 0.97 | 0.83 | 0.80 |
| P (g/l) | 24.2 | 15.5 | 13.2 | 11.6 | 10.8 |
| Y _{x/s} (g/g) | 0.22 | 0.217 | 0.212 | 0.192 | 0.187 |
| Y _{p/s} (g/l) | 4.662 | 3.209 | 2.88 | 2.672 | 2.529 |

Table 9

Effect of various nitrogen sources lactic acid production and biomass yield

| Kinetic parameters | Sodium nitrate (g/l) | Peptone (g/l) | Potassium nitrate (g/l) | Urea (g/l) | Yeast extract (g/l) |
|--------------------------|----------------------|---------------|-------------------------|------------|---------------------|
| S ₀ (g/l) | 18.38 | 17.82 | 18.855 | 18.3 | 17.78 |
| S (g/l) | 14.23 | 14.01 | 14.875 | 14.4 | 13.86 |
| S ₀ - S (g/l) | 4.15 | 3.81 | 3.98 | 3.87 | 3.92 |
| SC (%) | 22.57 | 21.38 | 21.10 | 21.2 | 22.04 |
| X (g/l) | 0.95 | 0.69 | 0.84 | 0.72 | 0.79 |
| P (g/l) | 9.4 | 5.4 | 8.2 | 5.6 | 7.4 |
| Y _{x/s} (g/g) | 0.228 | 0.181 | 0.211 | 0.19 | 0.20 |
| Y _{p/s} (g/l) | 2.265 | 1.417 | 2.06 | 1.44 | 1.887 |

Table 10

Effect of various concentrations of sodium nitrate

| Kinetic parameters | 6 (g/l) | 8 (g/l) | 10 (g/l) | 12 (g/l) | 14 (g/l) |
|--------------------|---------|---------|----------|----------|----------|
| S_0 (g/l) | 18.19 | 18.25 | 18.383 | 18.60 | 18.68 |
| S (g/l) | 14.23 | 14.23 | 14.23 | 14.79 | 14.98 |
| $S_0 - S$ (g/l) | 3.96 | 4.02 | .15 | 3.81 | 3.72 |
| SC (%) | 21.77 | 22.02 | 22.57 | 20.49 | 19.91 |
| X (g/l) | 0.86 | 0.90 | 0.95 | 0.83 | 0.76 |
| P (g/l) | 9.1 | 9.2 | 9.4 | 8.5 | 8.2 |
| $Y_{x/s}$ (g/g) | 0.217 | 0.223 | 0.229 | 0.218 | 0.204 |
| $Y_{p/s}$ (g/l) | 2.297 | 2.288 | 2.265 | 2.230 | 2.204 |

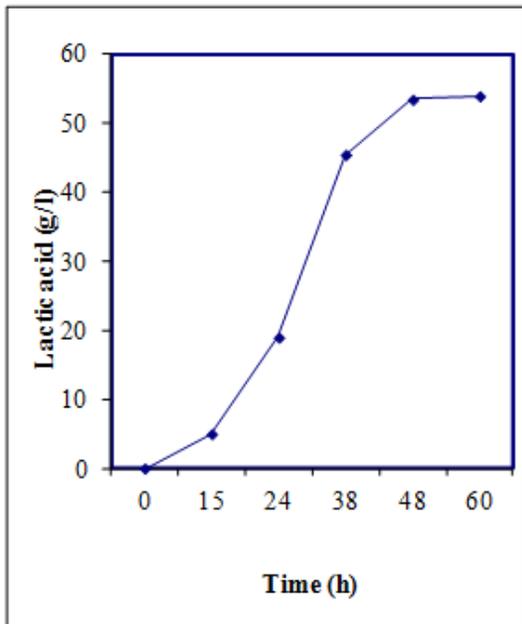


Fig.1: Effect of fermentation time on lactic acid production using *lactobacillus rhamnosus* MTCC1408 in whey + peptone (12 g/l) +mannitol(12 g/l)

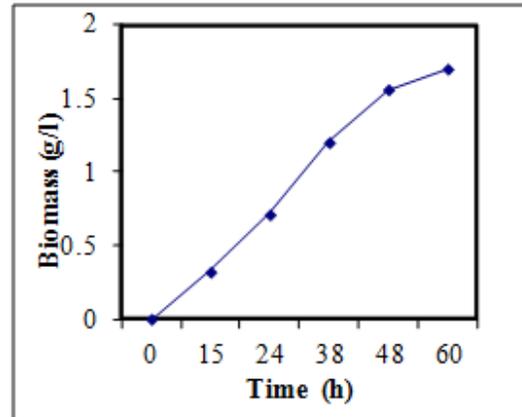


Fig. 2: Effect of fermentation time on biomass yield using *lactobacillus rhamnosus* MTCC1408 in whey + peptone (12 g/l) + mannitol(12 g/l)

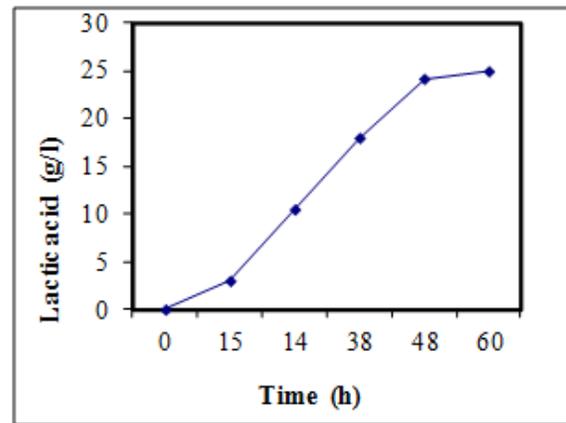


Fig. 3: Effect of fermentation time on lactic acid production using *lactococcus lactis subsp lactis* MTCC440 in whey + glucose (6 g/l) +sodium nitrate (10 g/l)

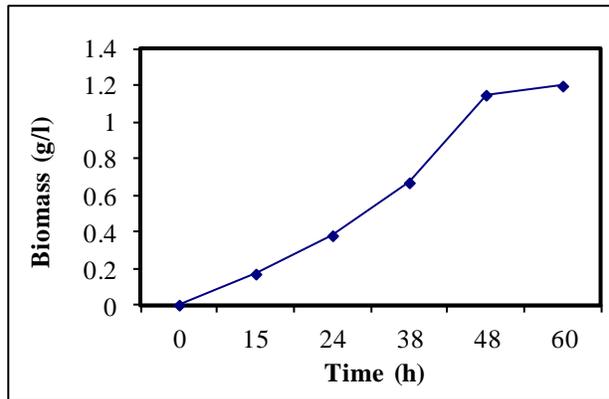


Fig. 4: Effect of fermentation time on biomass yield using *Lactococcus lactis subsp lactis* MTCC440 in whey + glucose (6 g/l) +sodium nitrate (10 g/l)

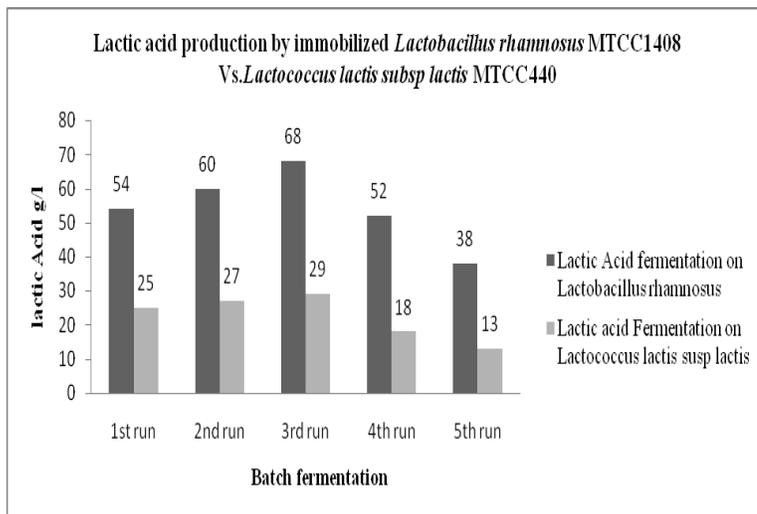


Fig. 5: Lactic acid production by immobilized *Lactobacillus rhamnosus* MTCC1408 and *Lactococcus lactis subsp lactis* MTCC440

REFERENCES

[1]. Saowanit Tongpim, Ratchanu Meoydong, Ayumi Naka and Kenji Sakai, Investigation of thermotolerant, lactic acid-producing bacteria from various sources in the North-east region of Thailand, Khon Kaen University

[2]. Kiremitci Gumusterelioglu and Gunday Teniz, Synthesis, Characterization and in vitro degradation of poly (DL - Lactide)/Poly (DL -Lactide - Co - Glycolide), *Turk J Che*, 23, 1999,153-161.

[3]. Xueliang Shen and Liming Xia, Lactic acid production from cellulosic waste by immobilized cells of *Lactobacillus delbrueckii*, *World J. Microbiol. Biotechnol.*, 22, 2006,1109-1114.

[4]. Junya Narita, Saori Nakahari, Hideki Fukuda and Akihiko Kondo, Efficient production of L- (+/-)Lactic acid from Raw Starch by *Streptococcus bovis* 148, *Journal of Bioscience and Bioengineering*, 97(6) , 2004,423-425.

[5]. Li Ping Huang, BO Jin, Paul Lant and Jiti Zhou, Simultaneous saccharification and fermentation of potato starch wastewater to lactic acid by *Rhizopus Oryzae* and *Rhizopus arrhizus*, *Biochemical Engineering Journal*, 23, 2005,265-276.

[6]. C.R. Soccol, V.I. Stonoga, and Raimbault, Production of L-Lactic acid by *Rhizopus* species, *World journal of Microbiology and Biotechnology*, 10, 1994, 433-435.

[7]. Vijai Pal, Marilingappa Jamuna and Kadirvelu Jeevaratnam, Isolation and characterization of bacteriocin producing lactic acid bacteria from a south indian special dosa (appam) batter, *Journal of Culture Collections*, 4, 2005, 53-60.

[8]. L .Levente, and Taya Puzanov, Membrane Fermentation of Lactic acid, *International Journal of Applied Science and engineering*, 3(1), 2005, 19-25.

[9]. Mohammad Altif, Basa Janakiram Nadeena and Gopal Reddy, Screening of inexpensive nitrogen sources for Production of L(+) Lactic acid from starch by amyolytic *Lactobacillus amylophilus* GV6 in single step fermentation, *Food Technol Biotechnol*, 43 (3), 2005,235-239.

[10]. A.Mona, M .Fayza and N.Mona, Lactic Acid Production and its Effect on Paper, *Journal Applied Sciences Researc.*, 2(11) 2006, 869-873.

[11]. Niju Narayanan, Pradip.K and Aradhana Srivaseapa, L(+) Lactic acid Fermentation and its Product Polymerization, *Electronic Journal of Biotechnolog.*, 7 (2), 2004, 165-234.

[12]. Yumiko Ohkouchi and Yuzo Inoue, Impact of chemical components of organic wastes on L(+)- Lactic Acid Production, *Resource Technology*, 98, 2007,546-553.

[13]. N. L. Tatsadjieu1, Y. N Njintang, T. Kembang Sonfack, B. Daoudou and C. M. F. Mbofu, Characterization of lactic acid bacteria producing bacteriocins against chicken *Salmonella enterica* and *Escherichia coli*, *African Journal of Microbiology Research* ,3 (5). 2009, 220-227.

Compression of Initial Ranging Scheme in WiMAX Using Markova Models

¹Dr.Balakrishna R, ²Vijayanand S, ³Rajesh K.S.

¹Professor & HOD, Dept of ISE, Rajarajeswari College of Engineering, Bangalore.
^{2,3}Asst.. Professor, Dept of ISE/CSE, Rajarajeswari College of Engineering, Bangalore

Abstract :

In this paper to analyze the Initial Ranging (IR) scheme in terms of the delay incurred and then proposes and implements an improvement to the same scheme. Initial Ranging is the process of obtaining the correct timing offsets and power adjustments such that the Subscriber Station (SS) is co-located with the Base Station (BS) and is transmitting at a suitable power level. During this process, which is a vital part of the network entry procedure in WiMAX, a series of request and response packets are exchanged between the Subscriber Stations and the Base Station. Due to the presence of multiple Subscriber Stations the request packets collide with each other thereby increasing the time required to complete the Initial Ranging process. Therefore, in our paper we analyze and obtain a formula for the delay incurred in Initial Ranging, using Markov models. In order to reduce the delay, we propose an enhancement to the existing Initial Ranging scheme. The concept of circularity is introduced, resulting in a reduced probability of collisions among the ranging request packets and compare the improved scheme with the original, and validate the delay formula obtained, using simulations.

Key words: Base station, Initial Ranging, WiMax,

Related work

Initial Ranging (IR) is the process of obtaining the correct timing offsets and power adjustments such that the Subscriber Station (SS) is co-located with the Base Station (BS) and is transmitting at a suitable power level. During this process, which is a vital part of the network entry procedure in WiMAX[1], a series of request and response packets are exchanged between the Subscriber stations and the Base Station. Due to the presence of multiple Subscriber stations the request packets collide with each other thereby increasing the time required to complete the Initial Ranging process. Therefore, in our project we analyze and obtain a formula for the delay incurred in Initial Ranging, using Markov models. Also, in order to reduce the delay involved, we propose an enhancement to the existing Initial Ranging scheme. The concept of circularity[2] is introduced to make the Subscriber stations less self-seeking,

resulting in a reduced probability of collisions among the ranging request packets. We then compare the improved scheme with the original, and validate the delay formula obtained, using simulations.

Introduction

Pervasive computing describes an environment where a wide variety of devices carry out information processing tasks on behalf of users by utilizing connectivity to wide variety of networks. Pervasive computing does not just mean "computers everywhere"; it means "computers, networks, applications, and services everywhere." Mark Weiser[3] was a pioneer of this field.

Pervasive computing is roughly the opposite of virtual reality. Where virtual reality puts people inside a computer-generated world, pervasive computing forces the computer to live out here in the world with people.

Pervasive computing creates an augmented reality. It enriches objects in the real world and makes them "smart." This allows these devices to better assist people. With additional information about the environment and the context, these devices become better tools for the people using them.

Pervasive Computing is also known as *Ubiquitous* or *Nomadic* computing

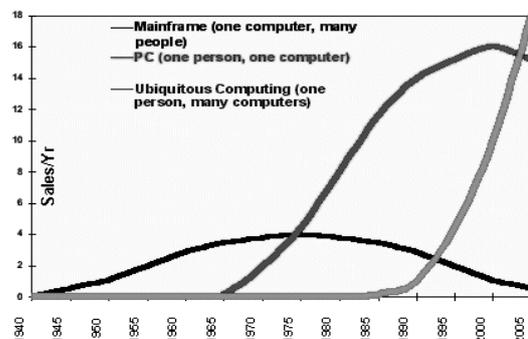


Figure 2.1: Major trends in Computing.

Ubiquitous[4] computing names the third wave in computing, just now beginning. First were mainframes, each shared by lots of people. Now we

are in the personal computing era, person and machine staring uneasily at each other across the desktop. Next comes ubiquitous computing, when technology recedes into the background of our lives.

1. Evolution of Pervasive Computing

Mobile computing and pervasive computing represent major evolutionary steps in a line of research dating back to the mid-1970s. Figure 2.2 illustrates this evolution from a systems-centric viewpoint. New problems are encountered as one move from left to right in this figure. In addition, the solutions of many previously-encountered problems become more complex - as the modulation symbols suggest, this increase in complexity is multiplicative rather than additive[4]. It is much more difficult to design and implement a mobile computing system than a distributed system of comparable robustness and maturity; a pervasive computing system is even more challenging. As Figure 2.3 indicates, the conceptual framework and algorithmic base of distributed systems provides a solid foundation for mobile and pervasive computing.

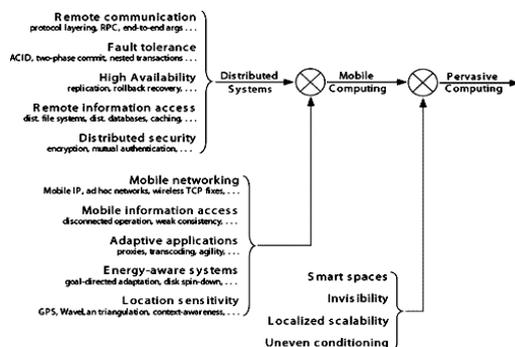


Figure 2.2: Evolution of Pervasive Computing
Characteristics of pervasive computing

Small: Ongoing miniaturization of components, moving to invisibility.

Embedded: Components are placed on or within other devices, objects or living beings.

Networked: Flexible capacity to exchange data and software components with other devices and platforms.

Context sensitive: Collect and exchange data on their environment and the host object via sensors.

Adaptive: Implement changes and modifications at the software and object level.

Collaborative: Ability to discover other objects and interact with them to establish cooperation on the software or information level

Network Volume: Sufficient in number and regularity of interaction to create network behaviors

This list of characteristics is interesting in how unremarkable it is. It indicates the extent to

which the trajectory of the technology used in pervasive computing has already been determined, even if the implementation remains problematic. In this sense, pervasive computing is not about the introduction of a single technology, but about a potential qualitative change that may arise through an increasingly integrated technological environment

In brief Pervasive Computing Model can be summarized as follows:

- Access control.
- Anytime.
- Anywhere.
- Any Device.
- Any Network.
- Any Data.

Coupled with intelligent applications the above model can be realized.

Access Control

Information available in any the Pervasive System should be protected and should be accessible only to authorized personal. Hence access control is a very important aspect of any Pervasive System.

Architecture of an access control mechanism:

- (1) Extract pieces of information in raw data streams early[5]
- (2) Define policies controlling access at the information level
- (3) Exploit information relationships for access control.

Anytime and Anywhere.

Pervasive Computing makes it possible users to access its services anytime and anywhere, round the clock.

Any Device

The power of Internet coupled with mobile devices such as PDAs, laptops envision the above goal.

The main features of pervasive devices are size, style, physical characteristics, content and services offered by them.

Based on size, there are three basic types of pervasive devices.

- Very Small Screen Devices (up to 4 inches)
- Small Screen Devices (up to 8 inches)
- Medium Screen Devices (up to 15 inches)

Any Network

Pervasive computing allows connectivity between different kinds of networks. Regardless of the network used at two ends, there can be a connection or information exchange between them.

Any Data

Data is classified as:

Categorical data- the objects being studied are grouped into some categories based on qualitative trait[6].

Measurement data- the objects being studied are grouped into some categories based on quantitative trait.

1. MARKOV MODELS

Markov processes provide very flexible, powerful, and efficient means for the description and analysis of dynamic (computer) system properties. Performance and dependability measures can be easily derived. Moreover, Markov processes constitute the fundamental theory underlying the concept of queuing systems. In fact, the notation of queuing systems has been viewed sometimes as a high-level specification technique for (a sub-class of) Markov processes[6].

Each queuing system can, in principle, be mapped onto an instance of a Markov process and then mathematically evaluated in terms of this process. But besides highlighting the computational relation between Markov processes and queuing systems, it is worthwhile pointing out also that fundamental properties of queuing systems are commonly proved in terms of the underlying Markov processes. This type of use of Markov processes is also possible even when queuing systems exhibit properties such as non exponential distributions that cannot be represented directly by discrete-state Markov models. Markovizing methods, such as embedding techniques or supplementary variables, can be used in such cases. Here Markov processes serve as a mere theoretical framework to prove the correctness of computational methods applied directly to the analysis of queuing systems. For the sake of efficiency, an explicit creation of the Markov process is preferably avoided.

Stochastic Process

A stochastic process is defined as a family of random variables $\{X_t : t \in T\}$ where each random variable X_t is indexed by parameter t belonging to T , which is usually called the time parameter if T is a subset of $R^+ = [0, \text{infinity})$. The set of all possible values of X_t (for each t belonging to T) is known as the state space S of the stochastic process[7].

If a countable, discrete-parameter set T is encountered, the stochastic process is called a discrete-parameter process and T is commonly represented by (a subset of) $N_0 = \{0, 1, \dots\}$; otherwise we call it a continuous-parameter process. The state space of the stochastic process may also be continuous or discrete. Generally, we restrict ourselves here to the investigation of discrete state spaces and in that case refer to the stochastic processes as chains, but both continuous- and discrete-parameter processes are considered[8].

A large number of stochastic processes belong to the important class of Markov processes. The theory of Markov chains and Markov processes is well established and furnishes powerful tools to solve practical problems. This chapter will be mainly devoted to the theory of discrete-time Markov chains, while the next chapter concentrates on continuous time Markov chains.

Definition of Markov process

A stochastic process $\{X(t), t \in T\}$ is a Markov process if the future state of the process only depends on the current state of the process and not on its past history. Formally, a stochastic process $\{X(t), t \in T\}$ is a continuous time Markov process if for all $t_0 < t_1 < \dots < t_{n+1}$ of the index set T and for any set $\{x_0, x_1, x_2, \dots, x_{n+1}\}$ of the state space it holds that

$$P_r[X_{k+1}=x_{k+1} | X_0=x_0, \dots, X_k=x_k] = P_r[X_{k+1}=x_{k+1} | X_k=x_k]$$

A Markov process is called a Markov chain if its state space is discrete. The conditional probabilities $P_r[X_{k+1} = j | X_k = i]$ are called the transition probabilities of the Markov chain. In general, these transition probabilities can depend on the (discrete) time k . A Markov chain is entirely defined by the transition probabilities and the initial distribution of the Markov chain $P_r[X_0 = x_0]$. Thus we obtain the following relations,

$$P_r[X_0=x_0, \dots, X_k=x_k] = P_r[X_k=x_k | X_0=x_0, \dots, X_{k-1}=x_{k-1}] * P_r[X_0=x_0, \dots, X_{k-1}=x_{k-1}]$$

By the definition of Markov model, we get the following relation,

$$P_r[X_0=x_0, \dots, X_k=x_k] = P_r[X_k=x_k | X_{k-1}=x_{k-1}] * P_r[X_0=x_0, \dots, X_{k-1}=x_{k-1}]$$

Discrete Time Markov Chain

If the transition probabilities are independent of time k ,

$$P_{ij} = P_r[X_{k+1} = j | X_k = i]$$

the Markov chain is called stationary. In the sequel, we will confine ourselves to stationary Markov chains. Since the discrete-time Markov chain is conceptually simpler than the continuous counterpart, we discuss with the discrete case.

4 .ANALYSIS AND DESIGN OF INITIAL RANGING

4.1 ANALYSIS OF IR SCHEME

After analyzing the Initial Ranging procedure, we enumerate the following states as well as transitions needed for modeling the procedure.

4.1.1 States involved

State 1: Waiting for UL-MAP. This is also the start state.

State 2: SS is performing Backoff procedure.

- State 3: Waiting for an RNG-RSP message from BS.
- State 4: Continue
- State 5: Success State – Wait for CDMA Allocation IE.
- State 6: Abort – Start network entry procedure at a different DL channel
- State 7: Waits for RNG-RSP again.
- State 8: Proceed to next phase of network entry
- State 9: Commence Periodic Ranging

4.1.2 Transitions involved

In State 1, the SS waits for a UL-MAP. After receiving this message it makes a transition to State 2. Transmission of CDMA code occurs at end of State 2. Also a timer is set for waiting for RNG-RSP message. This transition leaves the system in State 3.

When in State 3, if the timer for RNG-RSP expires then SS increments the power level and goes back to State 1.

When in State 3, if RNG – RSP is obtained with Ranging code as well as the Ranging slot, then it makes a transition to State 4. Here the necessary adjustments specified in RNG-RSP are made and system moves to State 1.

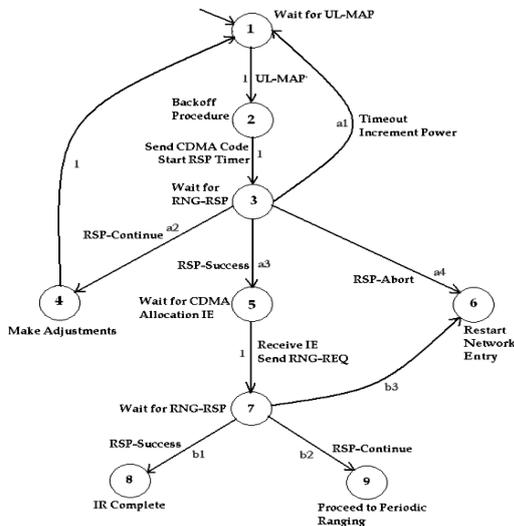


Figure 7.1 Markov Model of IR Scheme

When in State 3, if RNG-RSP is obtained with success status, then the system transits to State 5. Here it waits for CDMA Allocation IE. After reception it sends RNG-REQ message on the allocated bandwidth and moves to State 7[9].

When in State 7, on reception of RNG-RSP with success status it moves to State 8. On reception of RNG-RSP with continue status it moves to State 9. Else on reception of RNG-RSP with abort status, it goes to State 6 and SS starts the network entry procedure again.

When in State 3, if RNG-RSP is obtained with abort status then the system goes to State 6 and SS starts the network entry procedure again.

The following diagram shows the Markov model that represents the Initial Ranging procedure of IEEE 802.16 network standard.

The states 6, 8 and 9 lead out of the IR model and are the absorbing states.

Next, we use the transition matrix obtained above to obtain the overall delay formula. For this, we first need to tabulate the delays involved in the individual states[10].

4.1.3 Probabilities of Transition

The Transition Probability Matrix corresponding to the Markov model is as follows:

| | | | | | | | | | |
|---|----|---|---|----|----|----|---|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | a1 | 0 | 0 | a2 | a3 | a4 | 0 | 0 | 0 |
| 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | b3 | 0 | b1 | b2 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

Table 7.1 show transition probability matrix to the Markov model

The states 6, 8 and 9 are absorbing states. P (i, i) = 1.

| Delay Involved | Probabilities |
|---|---------------|
| UL-MAP Reception (1 to 2) | 1 |
| Backoff Delay + Sending CDMA (2 to 3) | 1 |
| RNG-RSP Timeout (3 to 1) | a1 |
| RNG-RSP Reception + Processing (3 to 4, 5 or 6) | a2,a3,a4 |
| IE Allocation Delay + Sending RNG-REQ (5 to 7) | 1 |
| RNG-RSP Reception + Processing (7 to 8, 9 or 6) | b1,b2,b3 |

Table 7.2 details of delay involved and probability

4.1.4 Details of Delays involved (Other than Backoff delay)

1. UL-MAP Reception = 5ms (Maximum of one complete frame length)
2. CDMA Sending Time = Transmission Time = 5ms/2 = 2.5ms [Frame Length/2 (Length of UL subframe) with frame length = 5ms]
3. RNG-RSP Timeout (T3) = 200 milliseconds.

4. RNG-RSP Reception + Processing (average value) = $T_3/2 + \text{Max. RNG-RSP Processing Time}/2 = 100 \text{ ms} + 10\text{ms}/2 = 105 \text{ ms}$
5. CDMA Allocation IE delay = 5s (same as 1)
6. Sending RNG-REQ (Same as 2) = 2.5ms
7. RNG-RSP Reception + Processing (average value) = 105ms

We assume that the delay involved for making changes at SS is negligible compared to the other delays involved.

4.1.5 Back-Off Delay Derivation

Consider the first time an SS enters Backoff procedure. Let the Initial Contention window be w_0 . The random number will be picked in the range $[0, w_0-1]$. Let this random number be called k. The SS has to defer a total of k contention slots. Let the number of CS's in a frame be n_{cs} . The number of frames that have to be deferred is k/n_{cs} . The delay involved here will be $(k/n_{cs}) * \text{frame length}$. After k/n_{cs} frames have passed the SS defers a further k modulus n_{cs} CS's. The delay involved here is equal to $(k \% n_{cs}) * T_{cs}$, where T_{cs} is the length of one CS [11].

Total delay so far = $(k/n_{cs}) * \text{frame length} + (k \% n_{cs}) * T_{cs}$

Here the value of k can vary from 0 to w_0-1 . Thus, we take an average of the delay over the random number.

$$AD_0 = (1/w_0) * \text{Sum of } [(k/n_{cs}) * \text{frame length} + (k \% n_{cs}) * T_{cs}] \text{ as k varies from 0 to } w_0 - 1.$$

Next we make an assumption that the probability of a successful transmission in a CS is 'p'. Thus, probability of failure will be '1-p'. In case of a failure the contention window is doubled in size. Let the new window be $[0, w_1-1]$. Similar to previous derivation the delay involved will be

$$AD_1 = (1/w_1) * \text{Sum of } [(k/n_{cs}) * \text{frame length} + (k \% n_{cs}) * T_{cs}] \text{ as k varies from 0 to } w_1 - 1.$$

1. Here $w_1 = 2 * w_0$.

Again there could be success or failure.

So, it will enter the third Backoff window phase $[0, w_2-1]$. Continuing in this fashion, we get the following delays for the next three phases.

$$AD_2 = (1/w_2) * \text{Sum of } [(k/n_{cs}) * \text{frame length} + (k \% n_{cs}) * T_{cs}] \text{ as k varies from 0 to } w_2 - 1.$$

$$AD_3 = (1/w_3) * \text{Sum of } [(k/n_{cs}) * \text{frame length} + (k \% n_{cs}) * T_{cs}] \text{ as k varies from 0 to } w_3 - 1.$$

$$AD_4 = (1/w_4) * \text{Sum of } [(k/n_{cs}) * \text{frame length} + (k \% n_{cs}) * T_{cs}] \text{ as k varies from 0 to } w_4 - 1.$$

Here $w_2 = 2 * w_1, w_3 = 2 * w_2, w_4 = 2 * w_3$.

We make another assumption at this point. The SS is assumed to complete successful transmission of its CDMA code, in a maximum of 5 Backoff phases. Thus, the worst case of transmission will be four failures followed by a success. The final formula for the delay will be as follows.

$$\begin{aligned} \text{Backoff Delay (BD)} &= p * \{AD_0 + t/2\} \\ &+ ((1-p)) * p * \{[AD_0 + t] + [AD_1 + t/2]\} \\ &+ ((1-p)^2) * p * \{[AD_0 + AD_1 + 2t] + [AD_2 + t/2]\} \\ &+ ((1-p)^3) * p * \{[AD_0 + AD_1 + AD_2 + 3t] + [AD_3 + t/2]\} \\ &+ ((1-p)^4) * p * \{[AD_0 + AD_1 + AD_2 + AD_3 + 4t] + [AD_4 + t/2]\} \end{aligned}$$

→ Equation 7.1: Backoff Delay incurred in IR scheme

Here t is the time-out after which failure is assumed. So, we take half that value for success i.e. $t/2$.

4.1.6 Overall Delay Formula

By traversing the transition diagram and multiplying the probabilities with the corresponding delays, the total delay can be calculated. The resulting formula is as follows.

The first part of the delay is in the loops 1-2-3-1 and 1-2-3-4-1. We call this D-loop. Then either success or abort occurs which is added to this part to get the final formula.

$$\begin{aligned} D_{loop} &= 1 * \text{UL-MAP} + 1 * (\text{BD} + \text{CDMA sending}) \\ &+ a_1 * (\text{Timeout } T_3 + \text{D-loop}) \\ &+ a_2 * (\text{RSP} + \text{D-loop}) \end{aligned}$$

Simplifying we get,

$$D_{loop} = \text{UL} + \text{BD} + \text{CDMA sending} + \frac{a_1 * T_3 + a_2 * \text{RSP}}{1 - (a_1 + a_2)}$$

→ Equation 7.2: The delay in the loops of IR Markov model

Now, the total delay involved can be represented using the formula given below.

$$\begin{aligned} D_{total} &= D_{loop} \\ &+ a_3 * (\text{RSP} + \text{CDMA_IE} + \text{RNG-REQ} + (b_1 + b_2 + b_3) * \text{RSP}) \\ &+ a_4 * \text{RSP} \text{ (here } b_1 + b_2 + b_3 = 1) \end{aligned}$$

→ Equation 7.3: Total delay in terms of loop delay

Substituting equation 2 in equation 3, we arrive at the final formula.

| |
|---|
| $D_{total} = \text{UL} + \text{BD} + \text{CDMA sending} + \frac{a_1 * T_3 + a_2 * \text{RSP}}{1 - (a_1 + a_2)} + a_3 * (\text{RSP} + \text{CDMA_IE} + \text{RNG-REQ} + \text{RSP})$ |
|---|

$$+ a_4 * RSP$$

→ Equation 7.4: Total delay incurred in the IR scheme

In equation 7.4 the Backoff Delay (BD) is given by equation 7.1

4.2 ENHANCEMENT OF IR SCHEME

During the IR procedure, the Subscriber Stations attempt to synchronize with the Base Stations. Since connections are not yet established, the process is contention based. As a result, there is always a finite probability that the request packets sent from different SS collide with each other. These packet collisions decrease the net throughput of the network, increase the Initial Ranging delay of SSs and ultimately lead to inefficient bandwidth utilization[12].

In our enhancement of the IR scheme, we attempt to reduce the probability of collisions between packets. We make use of the concept of circularity. Circularity is defined as a number which enables the identification of specific groups of events or packets. Each event or packet under consideration is numbered in a sequential manner. An event or packet with a number which is a multiple of the circularity value is said to be circularity-satisfied. It can be represented mathematically as follows

(Packet/Event number) = 0 modulo circularity

A finite delay is introduced before the occurrence of circularity satisfied events or the sending of circularity satisfied packets. This additional delay reduces the probability of packet collisions. The selfless behavior of certain SSs may increase the individual IR delays but on the whole the delay incurred in the entire network will be reduced.

In the IR procedure, whenever a ranging response packet is timed out, the backoff procedure is called and the window index is incremented. We apply the concept of circularity on this event. In case a time out event is circularity satisfied, the window index is increment an additional time. This, yields higher values for the backoff counter and hence reduces the packet collisions[5].

In addition, circularity is applied to ranging request packets in order to selectively increase the delay in their transmission. A finite delay is introduced before the circularity-satisfied request packets are sent. This further reduces request packet collisions.

5 . IMPLEMENTATION

The implementation consists of two parts:

- Simulation of existing scheme using ns-2
- Changing the backend and re-simulating

5.1 Simulation Setup

Tcl scripting is used to design and simulate the WiMAX[7] networks with varying architectures. Tcl gives us a lot of options that allow us to have a great degree of control over the

simulation of networks. Some of the important features of the Tcl script we have written are shown below.

5.1.1 Parameters Used

The following parameters are used for the simulation of the existing Initial Ranging scheme.

▪ General Parameters:

- Channel Type – WirelessChannel
- Radio Propagation Model– TwoRayGround
- Network Interface Type - Phy/WirelessPhy/OFDM
- MAC Type – 802_16
- Interface Queue Type – DropTail Priority Queue
- Link Layer Type – LL
- Antenna Model – Omni Antenna
- Maximum Packets in Interface Queue – 50
- Routing Protocol – DSDV (Routing is done through the Base Station)

▪ Network Architecture Parameters:

- Number of Base Stations – 1
- Number of Sink Nodes – 1
- Number of Subscriber Stations – Varied from 6 to 54
- Base Station Coverage – 20 meters
- Traffic Start Time – 20
- Traffic Stop Time – 40
- Simulation Stop Time – 50

5.1.2 Simulation of Existing Initial Ranging Scheme

The parameters mentioned are used in the Tool Command Language (TCL) script that we have written. This script also uses the WiMAX Control Agent in order to produce a detailed account of the activities going on during the simulations. In the resulting output file, we search for the timing details of specific events in order to extract the Initial Ranging delay.

We search for two events:

- 1) Found Ranging Opportunity – This marks the starting point of the IR procedure.
- 2) Ranging Response obtained with Success status – This marks the ending of the IR procedure.

The corresponding start and stop times of the IR procedure for all the Subscriber Stations (SS) in the scenario are stored in files. Using a C program, we find the average IR delay per node, after calculating the total time taken by all the nodes to complete their respective IR processes.

```

void RangingRequest::expire () {
    count1++;
    mac_->debug ("Ranging request expires\n");
    if(nb_retry_==(int)mac_-
>macmib_.contention_rmg_retry) {
//max retries reached, inform the scheduler
mac_->expire (type_);
} else {
if (window_ < s_->getBackoff_stop())
window_++;
nb_retry_++;
if (window_ < s_->getBackoff_stop() && count1 % 5
== 0)
window_++;
int result = Random::random() % ((int)(pow (2,
window_)+1));
mac_->debug ("Start Ranging contention in
%f(backoff=%d, size=%d, ps=%f)\n", result*s_-
>getSize()*mac_->getPhy()->getPS(),result,s_-
>getSize(),mac_->getPhy()->getPS());
    backoff_timer_->start (result*s_-
>getSize()*mac_-
>getPhy()->getPS());
        backoff_timer_->pause();
    }
}

```

Such simulations can be carried out for different numbers of SS each time by the use of shell scripts. Then the average IR delay is recorded along with number of SS involved in each such simulation[8].

5.2 Enhancement of Initial Ranging

In order to enhance the Initial Ranging scheme, we make some modifications to the backend of ns-2, which is implemented in C++ language. The files that are of interest to us are the following:

```

contentionrequest.cc
contentionslot.cc

```

During the IR procedure, there will be many SS contending to send their requests to join the network. The packets sent by different SS may collide at some instants and they will have to be resent. We try to reduce the collisions between packets of different SS by making the SSs less selfish. We have made two such changes, one in each of the files mentioned above. Let us see each one in detail[9].

5.2.1 Modification of 'contentionrequest.cc'

In this file, apart from many functions, the expire () function of the Ranging Request class is implemented. It contains the actions to be taken in case the ranging response from the BS is not received within a certain time limit. In the existing scheme, upon timing out the Backoff window index is incremented to the next valid number.

For the purpose of enhancement, we introduce a counter that is incremented every time the expire() function is called. Now we introduce a concept called circularity. Circularity is defined as

a number which enables the identification of specific groups of events or packets. It is implemented in terms of the modulo (%) operator. Consider the counter to be 'count1'. If 'count1' is a multiple of the circularity value, then the event is said to circularity satisfied. It can be represented mathematically as follows.

$$\text{count}_1 = 0 \text{ modulo circularity}$$

In case an 'expire' event is circularity satisfied, in order to reduce the probability of packet collisions, we increase the backoff window index one additional time. When we increase the backoff window an additional time, the backoff counter will assume larger values when chosen randomly. This reduces the chances of packet collisions from different SSs. After considering different circularity values, we observed that a value of 5 gives the maximum improvement in the delay. This is the first enhancement in our scheme. The code snippet below represents the improved 'expire()' function. The highlighted parts are the ones corresponding to the enhancement.

5.2.2 Modification of 'contentionslot.cc'

In this file, the addrequest() function of the RngContentionSlot class is implemented. This is a simple function in which, a new request packet is made ready in order to be sent during the ranging opportunity. The code snippet below shows the existing implementation.

As in the previous case, there is a certain finite probability of collision of these packets. In order to reduce the collision probability further, we keep track of the number of request packets being added. A global counter 'count2' is used that is incremented every time the addrequest() function is called at any SS. Now the principle of circularity is applied to these packets to identify the circularity-satisfied packets. The equation used is as follows.

$$\text{count}_2 \% \text{circularity} = 0$$

Those packets which have their counter value as multiples of the circularity value are delayed by a finite amount of time before they are sent out. This is achieved by using a call to the pauseTimers() function followed by a call to the resumeTimers() function. After considering different circularity values, we observed that a value of 3 gives the maximum improvement in the delay. The modified code is shown below with the added part highlighted.

```

void RngContentionSlot::addRequest (Packet *p)
{
    assert (request_ == NULL);
    count2++;
    request_ = new RangingRequest (this, p);
    if ( count2 % 3 == 0 ) {
        pauseTimers();
        resumeTimers();
    }
}

```

```
}
}
```

5.2.3 Simulation of Enhanced Initial Ranging Scheme

Since the backend of the ns-2 code is changed, we need to recompile the entire backend. After this is achieved, simulations are carried out in identical fashion to the simulations of original IR scheme.

6. RESULTS AND COMPARISON

CASE 1: For one base station and many subscribers stations (64) in which all the subscribers stations are placed at same positions.

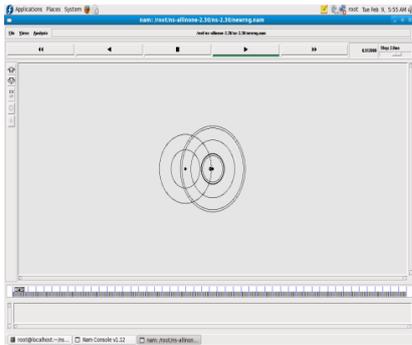


Fig 6.1 shows all mobile nodes access from same point

The delay values obtained for the above test case from the simulation are tabulated below.

| No of subscribers stations | Delay in milliseconds |
|----------------------------|-----------------------|
| 8 | 0.268365 |
| 16 | 0.454860 |
| 32 | 0.701238 |
| 64 | 0.845573 |

Table 6.1 Delay values for case 1

The graphical representation for the above tabulated delay values is shown below

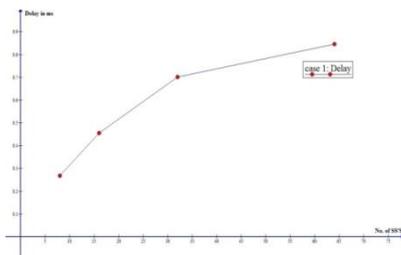


Fig 6.2 Graphical representation for delay values of case 1

The throughput values obtained for the above test case from the simulation are tabulated below.

| No of subscribers stations | Throughput in Mbps |
|----------------------------|--------------------|
| 8 | 0.48693 |
| 16 | 0.31706 |
| 32 | 0.23467 |
| 64 | 0.19960 |

Table 6.2 Throughput values for case 1

The graphical representation for the above tabulated throughput values is shown below

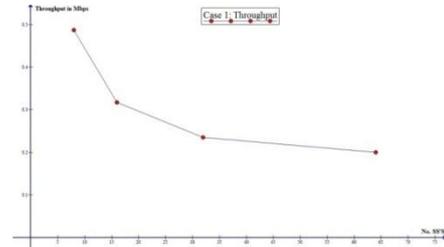


Fig 6.3 Graphical representation for throughput values of case 1

CASE 2: For one base station and many subscribers stations (64) in which all the subscribers stations are placed at circular positions around the base station.

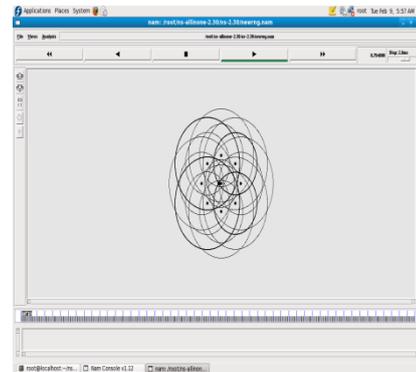


Fig 6.4 Shows mobile nodes access from circular position

The delay values obtained for the above test case from the simulation are tabulated below.

| No of subscribers stations | Delay in milliseconds |
|----------------------------|-----------------------|
| 8 | 0.2443542 |
| 16 | 0.4179602 |
| 32 | 0.5408570 |
| 64 | 0.6917536 |

Table 6.3 Delay values for case 2

The graphical representation for the above tabulated delay values is shown below

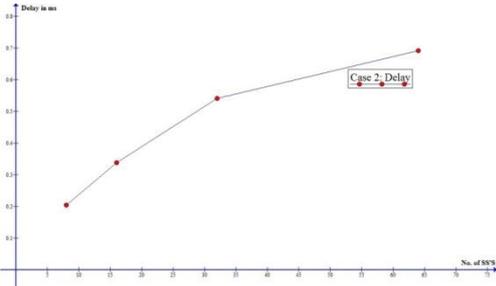


Fig 6.5 Graphical representation for delay values of case2

The throughput values obtained for the above test case from the simulation are tabulated below.

| No of subscribers stations | Throughput in Mbps |
|----------------------------|--------------------|
| 8 | 0.53665 |
| 16 | 0.38782 |
| 32 | 0.28228 |
| 64 | 0.23638 |

Table 6.4 Throughput values for case2

The graphical representation for the above tabulated throughput values is shown below

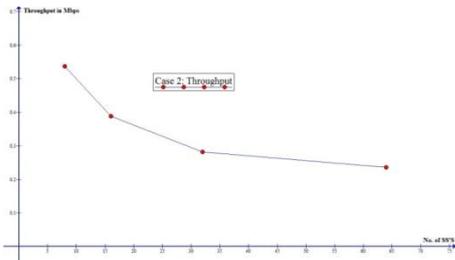


Fig 6.6 Graphical representation for throughput values of case2

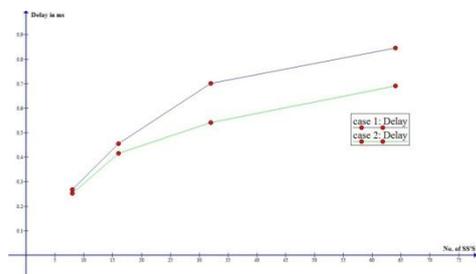


Fig 6.7 Comparison between case1 and case2 delay values

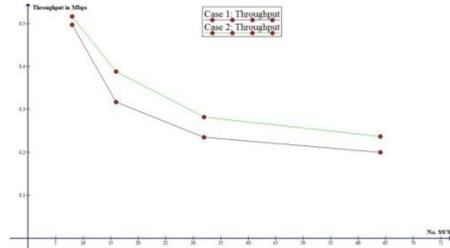


Fig 6.8 Comparison between case1 and case2 throughput values

CASE 3: For two base station and many subscriber stations (64) in which both the base stations have equal number of subscriber stations.

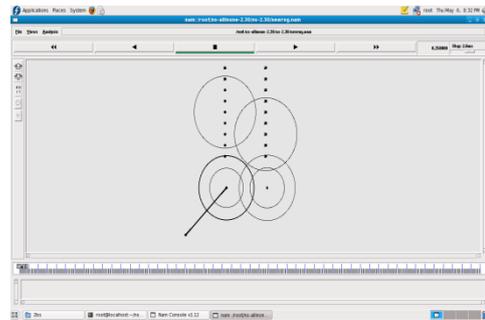


Fig 6.9 Shows 2BS having equal number of nodes

The delay values obtained for the above test case from the simulation are tabulated below.

| No of subscribers stations | Delay in milliseconds |
|----------------------------|-----------------------|
| 8 | 0.210553 |
| 16 | 0.404562 |
| 32 | 0.581445 |
| 64 | 0.765878 |

Table 6.5 Delay values for case3

The graphical representation for the above tabulated delay values is shown below

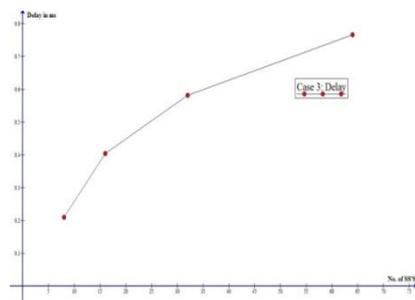


Fig 6.10 Graphical representation for delay values of case3

The throughput values obtained for the above test case from the simulation are tabulated below.

| No of subscribers stations | Throughput in Mbps |
|----------------------------|--------------------|
| 8 | 0.569927 |
| 16 | 0.296617 |
| 32 | 0.206382 |
| 64 | 0.156682 |

Table 6.6 Throughput values for case3

The graphical representation for the above tabulated throughput values is shown below

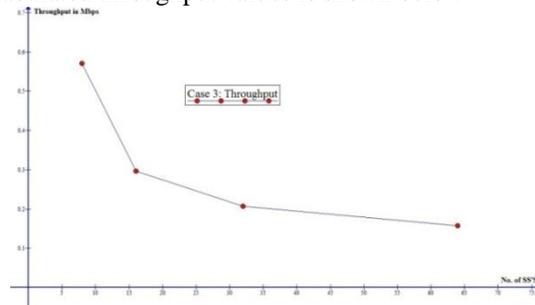


Fig 6.11 Graphical representation for throughput values of case3

These are the different cases on which we had successfully carried out simulation using NS-2. We had analyzed and evaluated delay and throughput for these above cases for IR mechanism. We had used various WiMAX patch files for its compatibility with NS-2. These patch files are backend files written in C++.

For the same purpose of analyzing and evaluating the performance of IR mechanism we had developed a generic code (shown in Appendix-B). This generic code written in visual C++ can be used to evaluate delay and throughput for further various cases.

CONCLUSION

In this paper, we have successfully analyzed and obtained a mathematical formula to calculate the delay involved in the Initial Ranging scheme. We have also enhanced this scheme using circularity, achieving about 25.10% reduction in the IR delay.

From the results and comparison we can also conclude that

1. With increasing number of SS, the circularity value controlling the delay must be increased.
2. With increasing number of SS, the circularity value controlling the window size must be decreased.

References

[1] *IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile*

Operation in Licensed Bands and Corrigendum 1, IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/ Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004) Std., 2006.

[2] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA system," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 296–306, Feb. 2004.

[3] X. Fu and H. Minn, "Initial uplink synchronization and power control (ranging process) for OFDMA systems," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 6, Nov. 29- Dec. 3, 2004, pp. 3999–4003.

[4] Jenhui Chen and Wei-Kuang Tan, "Predictive Dynamic Channel Allocation Scheme for Improving Power Saving and Mobility in BWA Networks," *ACM/Springer Mobile Networks and Applications (MONET)*, 2006.

[5] C. Eklund *et al.*, "IEEE Standard 802.16: A Technical Overview of the Wireless MAN Air Interface for Broadband Wireless Access," *IEEE Commun. Mag.*, 40(6):98–107, June 2002.

[6] B. Fong, N. Ansari, A.C.M. Fong, G.Y. Hong, and P.B. Rapajic, "On the Scalability of Fixed Broadband Wireless Access Network Deployment," *IEEE Commun. Mag.*, 42(9):12–21, September 2004.

[7] I. Koffman and V. Roman, "Broadband Wireless Access Solutions Based on OFDM Access in IEEE 802.16," *IEEE Commun. Mag.*, 40(4):96–103, April 2002.

[8] T. Kwon, et al., "Design and Implementation of a Simulator Based on a Cross-Layer Protocol between MAC and PHY Layers in a WiBro Compatible IEEE 802.16e OFDMA System," *IEEE Commun. Mag* 43(12):136–146, December 2005.

[9] Kyung-ah Kim, Chong-Kwon Kim, Tongsok Kim, A Seamless Handover Mechanism for IEEE802.16e Broadband Wireless Access, International Conference on Computational Science, vol. 2, pp. 527-534, 2005.

[10] Doo Hwan Lee, etc, "Fast Handover Algorithm for IEEE802.16e Broadband Wireless Access System", ISWPC, 2006

[11] Sik Choi, Gyung-Ho Hwang, Taesoo Kwon, "Fast Handover Scheme for Real-Time Downlink Services in IEEE 802.16e BWA System", Vol. 3 (2005), pp. 2028-2032, IEEE VTC 2005 spring, Sweden, May 2005

[12] Minsik Shim, Hwasung Kim, Sangho Lee, "A Fast Handover Mechanism For IPv6 Based WiBro System", Feb.20-22, 2006 ICAOT2006.

Implementation of Object Oriented Approach To Sequential Pattern Mining From Multidimensional Sequence Data

K.HEMALATA

2/2 M.TECH CSE, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

G.VASANTHAKUMARI

ASSOC. PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

Abstract— In recent years, emerging applications introduced new constraints for data mining methods. These constraints are typical of a new kind of data. Sequential pattern mining is applicable in this area, since many types of data sets are in a time related format. Besides mining sequential patterns in a single dimension, mining multidimensional sequential patterns can give us more informative and useful patterns. Due to the huge increase in data volume and also quite large search space, efficient solutions for finding patterns in multidimensional sequence data are nowadays very important. For this reason, in this paper, we present a multidimensional sequence model, Simulation experiments show good load balancing and scalable and acceptable speedup over different data sets and problem sizes.

Index Terms— Data Mining, Sequential Patterns, Sequence Data.

I. INTRODUCTION

Data mining has been defined as “The nontrivial extraction of implicit, previously unknown, and potentially useful information from data.” Mining frequent patterns or itemsets is a fundamental and essential problem in many data mining applications. These applications include the discovery of association rules, strong rules, correlations, sequential rules, episodes, multidimensional patterns, and many other important discovery tasks [1]. The architecture of a typical data mining system may have the following major Components:

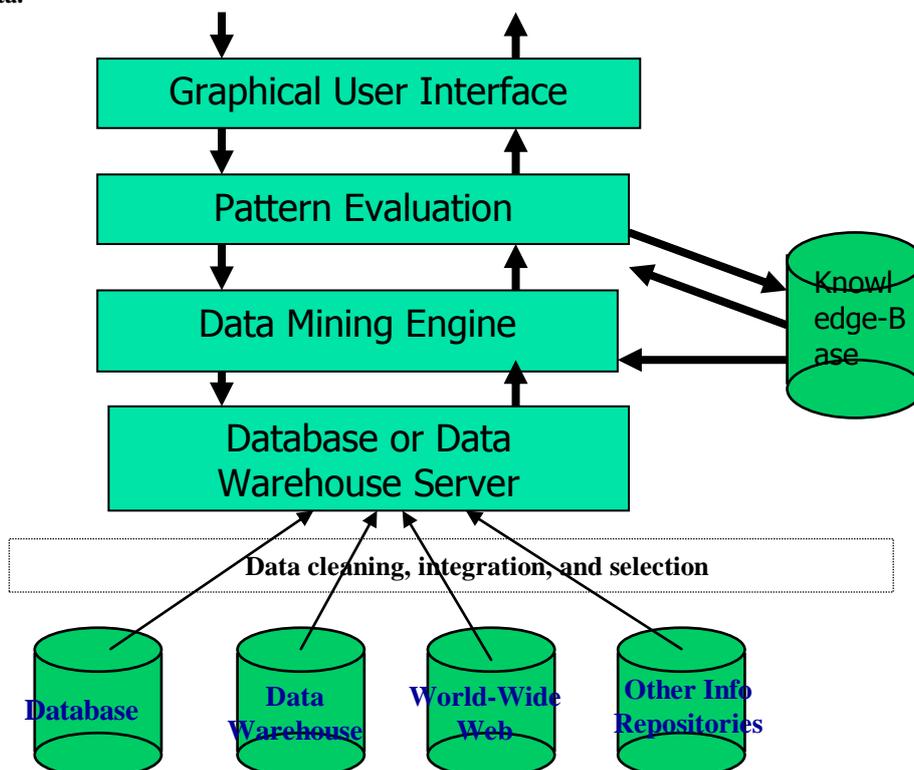


Fig. 1. Data mining Architecture

- **Database, Data warehouse, or other Information repository:** This is one or a set of databases, data warehouses, spreadsheets, or other kinds of information repositories. Data cleaning and data integration techniques may be performed on the data.
- **Database or Data warehouse server:** The database or data warehouse server is responsible for fetching the relevant data, based on the user's data mining request.
- **Knowledge base:** This is the domain knowledge that is used to guide the search, or evaluating the interestingness of resulting patterns. Such knowledge can include concept hierarchies, used to organize attribute or attribute values into different levels of abstraction.
- **Data mining engine:** This is essential to the data mining system and ideally consists of a set of modules for tasks such as characterization, association, classification, cluster analysis, and evolution and deviation analysis.
- **Pattern evaluation module:** This component typically employs interest measures and interacts with the data mining modules so as to focus the search towards interesting patterns.
- **Graphical user interface:** This module communicates between users and the data mining system, allowing the user to interact with the system by specifying a data mining query or task, providing information to help focus the search, performing exploratory data mining based on the intermediate data mining results.

One of the most important data mining problems is discovery of frequently occurring patterns in sequence data [2]. There are many domains where sequence mining has been applied, which include analysis of telecommunication systems, discovering customer buying patterns in retail stores, analysis of web access databases, and mining DNA sequences and gene structures.

Usually, sequence patterns are associated with different circumstances, and such circumstances form a multiple dimensional space. For example, bank customer sequences are associated with credit, branch, age, and others. It is interesting and useful to mine sequential patterns associated with multidimensional information [3].

Most of the proposed pattern mining algorithms are a variant of Apriori. Apriori-based algorithms show good performance with sparse datasets such as market-basket data, where the frequent patterns are very short. However, with dense data sets such as telecommunications, where there are many and long frequent patterns, performance of these algorithms degrades incredibly.

On the one side, parallel and distributed computing is expected to relieve current mining methods from the sequential bottleneck, providing the ability to scale to massive datasets, and improving the response time. Achieving good performance on

today's multiprocessor systems is a non-trivial task. The main challenges include synchronization and communication minimization, work-load balancing, find good data layout and data decomposition, and disk I/O minimization, which is especially important for data mining [4].

On the other side, the basic single dimension cannot satisfy the requirement of multi-attribute analysis, which is often the case in actual system practice. To address this problem, multidimensional sequence pattern mining is developed.

The most time consuming operation in the discovery process of sequential patterns is the computation of the frequency of the occurrences of interesting subsequences in the sequence database. However, the number of sequential patterns grows exponentially and the task of finding all sequential patterns requires a lot of computational resources, which make it an ideal candidate for parallel processing.

In this paper, we present a model for multidimensional sequence pattern and then propose mining sequential patterns from multidimensional sequence data. The experimental results show that our procedure usually achieves good load balancing and scalability.

II. SEQUENCE MINING AND RELATED WORK

The problem of mining frequent patterns in a set of data sequences together with a few mining algorithms was first introduced in [2]. They also presented three algorithms for solving this problem. The sequence mining task is to discover a sequence of attributes, shared across time among a large number of objects in a given database. Many studies have been contributed to the efficient mining of sequential patterns in the literature, most of which was focused on developing efficient algorithms for finding all sequential patterns such as , GSP [5], SPADE [6], Prefix Span [7] and so on. In addition, enormous sizes of available databases and possibly large number of mined sequential patterns demand efficient and scalable parallel algorithms.

Other work contributes on an extension of the problem of sequential pattern mining like constraint-based in sequential pattern mining, mining cyclically repeated patterns, approximate mining of consensus sequential patterns, mining top-k closed sequential patterns, mining frequent max sequential patterns, mining long sequential patterns in a noisy environment, mining hybrid sequential patterns and sequential rules, etc. Mining multidimensional sequential pattern is also one of the central topics in sequence mining as showed by the research efforts produced in recent years [8,9,10].

III. SYSTEM ARCHTECTURE

In this section we provide the step by step procedure for our proposed method, ask the input(file) based

on input file system will try to retrieve the data, consider the minimum support and confidence after that generate the frequent itemsets. After wards it will generate the support tree, it will a follow a set of

procedures finally generate association rules and finally display the items sets and AR rules as shown in the fig 2.

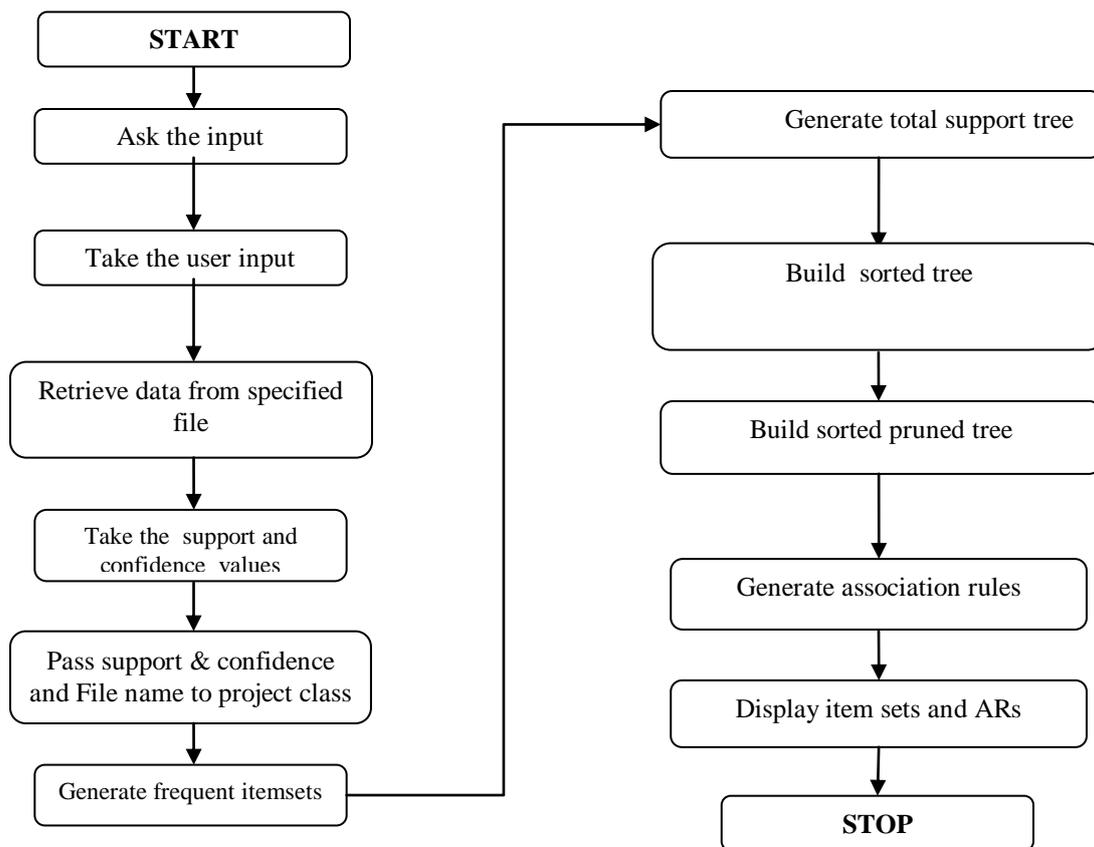


Fig 2 System Architecture

IV. MODULE DESCRIPTION AND IMPLEMENTATION

Design is the power to think, plan, and realize products that serve to accomplishment of any purpose. It is the process or art of defining the components, modules, interfaces, and data for a computer system to satisfy specified requirements. Here the system is divided into three modules.

A. Module1: Discovering the Apriori-T with X-Cheking.

To discover the patterns from the multi-dimensional sequence data. Each sequence pattern check with all possible sequence patterns based on the AprioriMD algorithm. First of all it will take only some records in the database according to the support and confidence levels. It will take the patterns in the tree levels and calculated no of records from each and every individual records and different combinational sequence patterns. It will display only the minimum threshold level.

B. Module2: Discovering Frequent Item Sets

This module deals to find out the most frequently accessed records or web pages. Customer has been visiting same sites in different sessions. From that each session it will calculate the no of frequent records or items. Each sequence pattern stored in

the database as tree structure. These sequence patterns kept in the levels of the tree.

C. Module3: Calculating the .Generate ARS

To generate the most frequently accessed records average to be calculated and give the percentage of used records. Using the all records in the database it will generate rules based on the support and confidence level. It has been pick up the records from above the support and confidence percentage.

V. IMPLEMENTATION

A platform is the hardware or software environment in which a program runs. Most platforms can be described as a combination of the operating system and hardware. The Java platform differs from most other platforms in that it's a software-only platform that runs on top of other hardware-based platforms. In this paper we generate the rules using an object oriented language called JAVA 1.6.

A. FREQUENT ITEMSET

Discovery of frequent item sets in transactional databases is a well-studied data mining topic. Given a database of transactions D and a collection of items L, an item set I belongs to L is frequent if the number of transactions in D that contain I (i.e., the support of I) is at least minimum support.

First, the input data set is scanned once to find all frequent 1-itemsets. These are sorted in descending order of their frequencies. Then, the data set is scanned again to construct an FP-tree. The FP-tree is a compact data structure that stores all patterns which appear in the database. For each transaction, the frequent items (in descending frequency order) are inserted as a path in this tree, where each node corresponds to an item and paths with common prefixes are compressed. An auxiliary structure, called header table, is used to link nodes with the same label. An algorithm, called FP-growth, is used to mine frequent patterns from the FP-tree.

B. AprioriMD:

AprioriMD, for mining the tree structure is a recursive procedure during which many sub trees and header tables are created. The process commences by examining each item in the header table, starting with the least frequent. For each entry the support value for the item is produced by following the links connecting all occurrences of the current item in the tree. If the item is adequately supported, then for each leaf node a set of *ancestor labels* is produced (stored in a *prefix tree*), each of which has a support equivalent to the sum of the leaf node items from which it is generated. If the set of ancestor labels is not null, a new tree is generated with the set of ancestor labels as the dataset, and the process repeated. In our implementation all frequent item sets thus discovered were placed in a T-tree, thus providing fast access during the final stage of the association rule mining (ARM) process, while at the same time providing for the deletion of sub trees and tables created "on route" as the AprioriMD algorithm progressed.

In AprioriMD, for mining sequential patterns from a multidimensional sequence database by modifying the Apriori or GSP algorithms. In the presentations afterward, we assume that the sequences are represented by simplified format. The algorithm is similar to the GSP algorithm. However, there are two differences:

- 1) The method to generate C_k and
- 2) The method to compute the supports of candidate sequences.

The following discusses how to handle these differences. We first discuss how to generate C_2 .

Then, we discuss how to generate C_k , where $k > 2$. Traditionally, C_2 can be obtained by joining L_1 with L_1 directly. However, since the first element and the second element in C_2 , say b and c , may have different dimensional scopes, we need to generate the pairs for all possible scope relations.

C. ASSOCIATION RULE MINING

The T-tree (Total Support Tree) is a compressed set enumeration tree structure. The T-tree differs from more standard set enumeration trees in that the nodes at the same level in any sub-branch are organized into 1-D arrays such that the array indexes represent column numbers.

For this purpose it is more convenient to build a "reverse" version of which permits direct indexing with attribute/column numbers. The T-tree offers two initial advantages

1. Fast traversal of the tree using indexing mechanisms,
2. Reduced storage, in that item set labels are not required to be explicitly stored, and thus no sibling reference variables (pointers) are required.

As each level is processed, candidates are added as a new level of the T-tree, their support is counted, and those that do not reach the required threshold of support are subsequently pruned. When the algorithm terminates, the T-tree contains only the large item sets. At each level, new candidate K item sets are generated from identified large $K-1$ item sets, using the *downward closure property of item sets*, which in turn may necessitate the inspection of neighboring branches in the T-tree to determine if a particular $K-1$ subset is supported. We refer to this process as *X-checking*. Note that X-checking adds a computational overhead; offset against the additional effort required to establish whether a candidate K item set, all of whose $K-1$ item sets may not necessarily be supported, is or is not a large item sets. In some cases it is more expedient to assume that those subsets of a candidate K item sets that are contained in neighboring branches of the T-tree are supported than to carry out X-checking. Results are shown in the Fig 3,4,5.

```
C:\WINDOWS\system32\cmd.exe

D:\mallik\seq>java AprioriTapp -Finput.txt -S50 -C50
SETTINGS
-----
File name           = input.txt
Support (default 20%) = 50.0
Confidence (default 80%) = 50.0

Reading input file: input.txt
Number of records = 768
Number of columns = 42
Min support       = 384.0 (records)
APRIORI-T WITH X-CHECKING
-----
Minimum support threshold = 50.0% (384.0 records)
Levels in T-tree = 3
Generation time = 0.02 seconds (0.0 mins)
[1] {1} = 424
[2] {21} = 642
[3] {28} = 492
[3.1] {28 21} = 404
[4] {31} = 524
[4.1] {31 21} = 457
[5] {36} = 457
[5.1] {36 21} = 385
```

Fig 3 Input Screen

```
C:\WINDOWS\system32\cmd.exe

APRIORI-T WITH X-CHECKING
-----
Minimum support threshold = 50.0% (384.0 records)
Levels in T-tree = 3
Generation time = 0.02 seconds (0.0 mins)
[1] {1} = 424
[2] {21} = 642
[3] {28} = 492
[3.1] {28 21} = 404
[4] {31} = 524
[4.1] {31 21} = 457
[5] {36} = 457
[5.1] {36 21} = 385
[6] {41} = 500
[6.1] {41 21} = 439
Number of frequent sets = 10
Number of Nodes created = 57
Number of Updates       = 12084
T-tree Storage         = 1132 (Bytes)
```

Fig 4 Frequent sets displayed screen

```

C:\WINDOWS\system32\cmd.exe
PREQUENT <LARGE> ITEM SETS:
-----
[1] {1} = 424
[2] {21} = 642
[3] {28} = 492
[4] {28 21} = 404
[5] {31} = 524
[6] {31 21} = 457
[7] {36} = 457
[8] {36 21} = 385
[9] {41} = 500
[10] {41 21} = 439

GENERATE ARs :
-----
(1) {41} -> {21} 87.8%
(2) {31} -> {21} 87.21%
(3) {36} -> {21} 84.24%
(4) {28} -> {21} 82.11%
(5) {21} -> {31} 71.18%
(6) {21} -> {41} 68.38%
(7) {21} -> {28} 62.92%
(8) {21} -> {36} 59.96%

D:\mallik\seq>_

```

Fig 4 AR rules displayed screen

VI. CONCLUSIONS AND FUTURE WORK

Multidimensional mining has been attracting attention in recent research into data mining. Very large search space and data volume have made many problems for mine sequential patterns. In order to effectively mine, efficient parallel algorithm is necessary. In this paper, we theoretically present a multidimensional sequence model and then use this strategy to multidimensional sequential pattern mining.

In this paper we developed a modification of the well-known Apriori algorithm for mining sequential patterns from multidimensional sequence data. The simulation results show that the latter is better than the former.

Future work

Scalability is important for any system; various combinations of algorithms may be used for achieving better result. The accuracy of the mining may be improved by preprocess the data before analysis. So in this way there is scope to the future enhancements.

REFERENCES

- [1] J. Han, and M. Kamber, Data Mining: Concepts and Techniques, Academic Press, New York, 2001.
- [2] R. Agrawal, R. Srikant, Mining Sequential Patterns, Proc. of the 11th Int'l Conference on Data Engineering, 1995.
- [3] J. Han, J. Pei, K. Wang, Q. Chen, and U. Dayal, Multi-dimensional sequential pattern mining, Conference on Information and Knowledge Management, pp: 81-88, 2001 .
- [4] M. J. Zaki, C.-T. Ho, Large scale parallel data mining, Lecture notes in computer science, Vol 1759: Lecture notes in artificial intelligence, Springer, 2000 .
- [5] R. Srikant, R. Agrawal, Mining sequential patterns: generalizations and performance improvements, In Proc. of the 5th Int'l Conference on Extending Database Technology, pp: 3-17, 1996 .
- [6] M. J. Zaki, An efficient algorithm for mining frequent sequences, Machine Learning Journal, vol. 42(1/2), pp: 31-60, 2001 .
- [7] J. Pei, J. Han, H. Pinto, Q. Chen, U. Dayal, MC. Hsu, PrefixSpan: Mining sequential patterns efficiently by prefix-projected pattern growth, In Proc. of Int'l Conference on Data Engineering, pp: 215-224, 2001.
- [8] S. de Amo, D. A. Furtado, A. Giacometti, and D. Laurent, An apriori-based approach for first-order temporal pattern mining, In Simposio Brasileiro de Bancos de Dados, 2004.

- [9] Plantevit M., Choong Y.W., Laurent A., Laurent D., Teisseire M. , M2SP: Mining Sequential Patterns Among Several Dimensions, Principles of Knowledge Discovery in Databases, PKDD, Volume 3721, pp: 205-216, 2005 .
- [10] C.-C. Yu, and Y.-L. Chen, Mining Sequential Patterns from Multidimensional Sequence Data, IEEE Transactions on Knowledge and Data Engineering archive, Volume 17, Issue 1, pp: 136-140, 2005.

Implementation of Network simulation for Distributed Explicit Rate Schemes in Multi-I/O Network Systems

DEVI PRIYA G

2/2 M.TECH CSE, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

B. RAMESH NAIDU

ASSOCIATE PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

Abstract— This paper describes a novel multi-rate multicast congestion control scheme based on the well-known proportional plus integrative control technique, where the control parameters can be designed to ensure the stability of the control loop in terms of source rate. The congestion controller is located at the next upstream nodes of multicast receivers and has explicit rate (ER) algorithm to regulate the rate of the receivers. We further analyze the theoretical aspects of the proposed algorithm, show how the control mechanism can be used to design a controller to support many-to-many multi-rate multicast transmission based on ER feedback, and verify its agreement with simulations in the case of bottleneck link appearing in a multicast tree. Simulation results show the efficiency of our scheme in terms of the system stability, high link utilizations, fast response, scalability, high throughput and fairness.

Index Terms— explicit rate, multicast congestion control, multi-rate multicast, QoS (quality of service), rate-based congestion control.

I. INTRODUCTION

With the ever-increasing wireless multicast data applications recently, considerable efforts have focused on the design of flow control schemes for Multicast congestion avoidance. There are generally two types of wireless multicast rate control schemes: Single-Rate Multicast (SR-M) and Multi-Rate Multicast (MR-M) [1, 2]. The SR-M is not fair to those receivers who are connected to high speed networks and are able to receive data at higher rates. Due to the diverse characteristics and requirements of receivers within a multicast group, it is desirable to have multicast sessions in which different receivers receive data at different rates. This is achieved by MR-M, where the source is able to transmit data to all receivers at different rates that suits the capacity of each individual receiver. Since in MR-M the capacities of network links to different

receivers differ and traffic should be accordingly adjusted at the links with different capacities, flow control becomes a very challenging issue. For simplicity, we use multicast to refer wireless MR-M for the rest of the paper, unless otherwise specified.

Several multicast congestion control approaches [3,4] have been proposed recently. One class of approaches adopts a simple hop-by-hop feedback mechanism. Although the simple hop-by-hop feature seems to be an advantage, these approaches often lead to the so-called consolidation noise problem [5, 6] due to incomplete feedback information. To overcome this drawback, Xiong et.al [7] proposed the concept of feedback synchronization, at each branch point, by accumulating feedback from all downstream branches. These schemes of [6] and [7] then introduce another problem of slow transient response since the feedback from the congested branch may have to needlessly wait for the feedback from "longer" paths. Such delayed congestion feedback can cause excessive queue build-up and packet loss at the bottleneck link. The authors of [8] and [9] suggested that only a carefully chosen set of receivers, instead of all receivers, send their feedbacks to the sender. Zhang et al. [10] proposed an optimal second-order rate control algorithm to deal with control packet round-trip time (RTT) variation in multicast communications, which defined that the data transfer rate is adjusted at the source depending on the available bandwidth at the bottleneck. More recently, several studies (such as [11, 12]) have focused on the design of MR-MCC protocols. However, all of them have drawbacks. Some designs cause over-subscription and high packet losses. Some are slow to converge and unresponsive. Some designs are too complex and infeasible [13]

II. RELATED WORK

The advances in multi-input–multi-output (MIMO) systems and networking technologies introduced a revolution recently, which promises significant impact in our lives. Especially with ever-increasing multicast data applications, wireless and wired multicast (multipoint-to-multipoint) transmission has considerable effect on many applications such as teleconferencing and information dissemination services. Multicast improves the efficiency of multipoint data distribution from multiple sender's to a set of receivers [14,15].

This paper describes a novel MR-MCC congestion control scheme based on the proportional plus integrative (PI) controller. The incoming flow rate of a session, at every branching point in its tree, is enforced to be the maximum of the rates that can be accommodated by its participating branches. By doing so, the sending rate at the source will eventually be the maximum of the rates that can be accommodated by the entire paths to individual receivers. Since the source sends data at the maximum path rate, it is necessary to reduce the rate of an incoming flow at every branching point to the value that can be accommodated by its participating branches [13]. The PI controllers are located at the next upstream branch node of the receivers.

The relevant gain parameters of the PI controller are determined by the system stability. Each branch point in our scheme only receives feedbacks from the direct downstream nodes instead of all downstream nodes, thus it greatly decreases the number of feedbacks to be aggregated at one node. As a result, our scheme can avoid the so-called feedback explosion problem [13] to a great

extent. Simulation results show the efficiency of the proposed scheme in terms of system stability, high link utilizations, quickly response, scalability, high system transport rate, intra-session fairness and intersession fairness. Simulation results verify the efficiency of the proposed MR-MCC scheme. Our scheme is very versatile. It can support sessions where receivers are added and depart. It can manage the traffic to guarantee stability, in real time, even if considerable changes occur in the source-receivers tree.

III. PROPOSED SYSTEM ARCHITECTURE

To analyze the performance and characteristic of the multicast, we focus on the following system model as shown in figure 1, where we have two classes of sources, i.e., one multicast source and one end-to-end CBR source. The PI controllers are located at the next upstream nodes of the receivers, i.e., the routers from RT_1 to RT_m , and compute the expected rates used to adjust the multicast receiving rates of the downstream receivers. The receiver j_i represents the i th receiver corresponding to the j th router (RT_j). We provide rate adaptation functionality at every branch point of each session. This rate adaptation scheme is determined on the basis on the fact that the multicast tree will eventually receive data at an independently trimmed rate allowed by its entire path. So we acquire the above computed maximum value as the effective sending rate of the multicast source. The sending rate is necessary to convert down the rate of an incoming flow at every branching point to the values that can be accommodated by its participating branches to individual receiver.

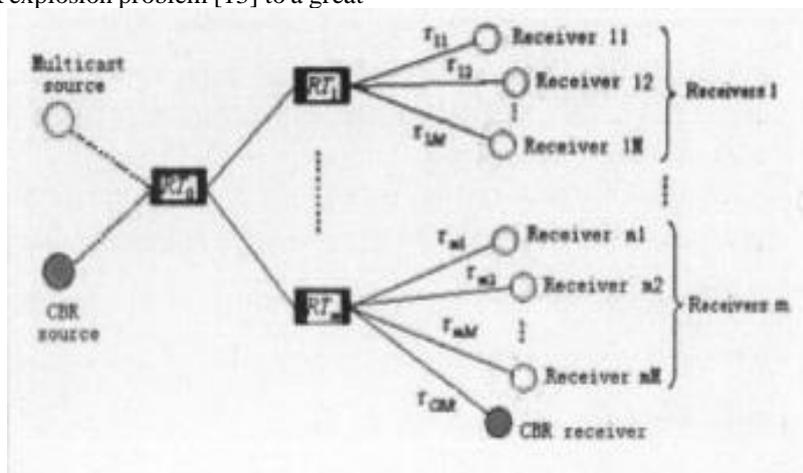


Fig 1 System Architecture

Algorithm

| |
|--|
| <p>Source Algorithm</p> <p>Upon every T epoch Transmit data including FCP; Upon multicast source receives a consolidation BCP from its downstream Adjust the transmitting rates in terms of min(the maximum receiving rate of corresponding receivers in the consolidated BCP, the bandwidth of the connective link);</p> |
| <p>Router Algorithm</p> <pre> If multicasttree[i]==1 { If the packet is an FCP { Put the packet in the buffer; Multicast the data packet including FCP to the downstream nodes; } else { If the node is the next upstream node of the receiver j { Computer the expected sending rate R_j for the receivers i using PI cotroller; } else { Select the maximum expected incoming rate of the next downstream node; } Construct the BCP based on the received BCP's and the relevant case; Feedback it to the upstream node; If receivedtree[i]==1 { Delete the data packets from the buffer; } Else { Maintain the data packets in the buffer until received all confirmations of the receivers; } } } </pre> |
| <p>Receiver Node Algorithm</p> <p>Upon receipt of an FCP Put the data packets into the buffer; Construct the BCP based on the current case of the receiver nodes; Feedback the BCP to the upstream branch point;</p> |

A. Merit of proposed system:

1. data transfer rate is adjusted at the source
2. group node makes sure that the buffer occupancy stabilizes and never overflows the buffer capacity.
3. these are active and effective methods to adjust the different sending rates to different receivers, and reduce the packets loss.
4. the main proposed scheme in terms of system stability and fast response to the buffer occupancy, as well as controlled sending rates, low packet loss, and high scalability.

IV. MODULE DESCRIPTION

Design is the power to think, plan, and realize products that serve to accomplishment of any purpose. It is the process or art of defining the components, modules, interfaces, and data for a computer system to satisfy specified requirements. Here the system is divided into three modules.

- Multicast Network Configuration Module:
- Multi-rate-multicast control (MR-MCC) tree Module:
- PI controllers Module:

A. Multicast Network Configuration Module:

The multicast network is a connection-oriented one, which is composed of sources and destination nodes. Multicast connection and every sampling period, the multicast source issues and transmits a FCP to the downstream nodes (the branch node and destination nodes), and a BCP is constructed by each branch node based on the PI controller because PI is located at each branch point, and sent back to the source. After the multicast source receives the BCPs from the downstream nodes, it will take appropriate action to adjust its transmitting rates of multicast traffic based on the computed value of the PI controller. After receiving the data packets coming from the branch point, the receivers construct BCPs and send them back to the branch point.

B. Multi rate-multicast control (MR-MCC) tree Module: users who are willing to pay more to access at a higher speed. Furthermore, due to the diverse characteristics and requirements of the different receivers within a multicast group, and for greater flexibility in resource allocation, it is desirable to have multicast sessions in which different receivers receive data at different rates. This inflexibility is overcome by MR-MCC that can allocate different rates.

C. PI controller Module: The PI controllers are located at the next upstream branch node of the receivers. The relevant gain parameters of the PI controller are determined by the system stability. Each branch point in our scheme only receives feedbacks from the direct downstream nodes instead of all downstream nodes, thus it greatly decreases the number of feedbacks to be aggregated at one node. As a result, our scheme can avoid the so-called feedback explosion problem [24] to a great extent. The incoming flow rate of a session, at every branching point in its tree, is enforced to be the maximum of the rates that can be accommodated by its participating branches. By doing so, the sending rate at the source will eventually be the maximum of the rates that can be accommodated by the entire paths to individual receivers.

First we select the source info which is required to transfer and based on our requirement we have to select the groups. At the receiving end the concerned receiving party receives the file. Simulation results show that the proposed approach decreased time, increases the through put and performance wise it is better compared to other schemes.

V. PERFORMANCE EVALUATION

Here we pay more attention to sending multi-rates of sources, buffer occupancy, link utilization, receiving rates of routers and end-users. We assume that the link delay is dominant compared to the processing delay or queuing delay.

Simulation Model The simulation model is shown in Fig. 2. There are different receivers, groups group1, group2,etc.,For convenience, we group together the receivers having similar receiving rates. Thus, we select a single receiver in each group as a representative of the group.

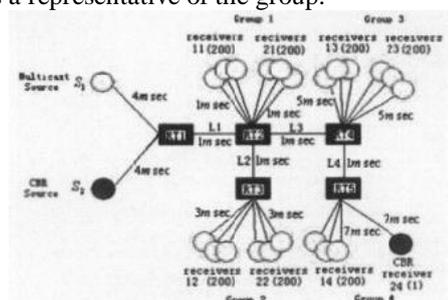


Figure2: The Simulation model

VI. CONCLUSIONS AND FUTURE SCOPE

This paper presents a theoretic analysis and design method of MR-MCC using explicit rate feedback mechanism to satisfy the different needs of the multiple users. The PI controller, whose control parameters can be designed to ensure the stability of the control loop in terms of buffer occupancy on the basis of control theory, is used in the next upstream node of the receivers to regulate the receiving rate. Relevant pseudo codes for implementation have subsequently been developed. It is clearly that the proposed MR-MCC scheme solves intra-protocol unfairness and low link utilization of SRMCC. Simulations have been carried out with a multicast source and a CBR source. Simulation results demonstrate the efficiency of our scheme in terms of the system stability, high link utilizations, fast response, scalability, high unitary throughput, intra-session fairness and inter-session fairness.

Future scope

Scalability is important for any system; various combinations of algorithms may be used for achieving better result. So in this way there is scope to the future enhancements.

REFERENCES

- [1] H. W. Lee and J. W. Cho. A distributed max-min flow control algorithm for multi-rate multicast flows. In Proceedings of IEEE Global Communications Conference, pages 1140-1146, vol. 2, Dallas, Texas, USA, November 29-December 3, 2004.
- [2] S. Sarkar and L. Tassiulas. Fair distributed congestion control in multirate multicast networks. IEEE/ACM Transactions on Networking, 13(1):121-133, February 2005.
- [3] L. Benmohamed and S. M. Meekov, "Feedback control of congestion in packet switching networks: The case of a single congested node," IEEE/ACM Trans. Netw., vol. 1, no. 6, pp. 693-708, Dec. 1993.
- [4] N. X. Xiong, Y. He, L. T. Yang, and Y. Yang, "A self-tuning reliable dynamic scheme for multicast flow control," in Proc. 3rd Int. Conf.

- Ubiquitous Intell. Comput. (UIC 2006), Wuhan, China, Sep. 3–6, pp. 351–360.
- [5] L. Lao, J. H. Cui, M. Gerla, and D. Maggiorini, “A comparative study of multicast protocols: Top, bottom, or in the middle,” presented at the IEEE INFOCOM Conf., Barcelona, Spain, Apr. 23–29, 2006.
- [6] A. Kolarov and G. Ramamurthy, “A control theoretic approach to the design of an explicit rate controller for ABR service,” *IEEE/ACM Trans. Netw.*, vol. 7, no. 5, pp. 741–753, Oct. 1999.
- [7] N. X. Xiong, Y. He, L. T. Yang, and Y. Yang, “A self-tuning reliable dynamic scheme for multicast flow control,” in Proc. 3rd Int. Conf. Ubiquitous Intell. Comput. (UIC 2006), Wuhan, China, Sep. 3–6, pp. 351–360.
- [8] F. Blanchini, R. L. Cigno, and R. Tempo, “Robust rate control for integrated services packet networks,” *IEEE/ACM Trans. Netw.*, vol. 10, no. 5, pp. 644–652, Oct. 2002.
- [9] R. Jain, S. Kalyanaraman, R. Goyal, S. Fahmy, and R. Viswanathan, “ERICA switch algorithm: A complete description,” in ATM Forum-TM, Aug. 1996, pp. 96–1172.
- [10] S. H. Lee and J. T. Lim, “Multicast ABR service in ATM networks using a fuzzy-logic-based consolidation algorithm,” *Proc. Inst. Electr. Eng. Commun.*, vol. 148, no. 1, pp. 8–13, Feb. 2001.
- [11] X. Zhang, K. G. Shin, D. Saha, and D. D. Kandlur, “Scalable flow control for multicast ABR services in ATM networks,” *IEEE/ACM Trans. Netw.*, vol. 10, no. 1, pp. 67–85, Feb. 2002.
- [12] N. Xiong, Y. He, Y. Yang, L. T. Yang, and C. Peng, “A self-tuning multicast flow control scheme based on autonomic technology,” in Proc. 2nd IEEE Int. Symp. Dependable, Auton. Secure Comput. (DASC), Indiana Univ., Purdue Univ., Indianapolis, IN, Sep. 29–Oct. 1, 2006, pp. 219–226.
- [13] J. Widmer and M. Handley, “Extending equation-based congestion control to multicast applications,” in Proc. ACM SIGCOMM, San Diego, CA, Aug. 27–31, 2001, pp. 275–286.
- [14] R. H. Gau, Z. J. Haas, and B. Krishnamachari, “On multicast flow control for heterogeneous receivers,” *IEEE/ACM Trans. Netw.*, vol. 10, no. 1, pp. 86–101, Feb. 2002.
- [15] L. Lao, J. H. Cui, M. Gerla, and D. Maggiorini, “A comparative study of multicast protocols: Top, bottom, or in the middle,” presented at the IEEE INFOCOM Conf., Barcelona, Spain, Apr. 23–29, 2006.

Implementation Of Object Oriented Approach To Authentication Of Group Key Transfer Using Secret Sharing

D. JEEVARATNAM

2/2 M.TECH CSE, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
ST.THERESSA INSTITUTE OF ENGINEERING TECHNOLOGY,
GARIVIDI, ANDHRA PRADESH, INDIA

K. VISALA

ASSISTANT PROFESSOR,
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
ST.THERESSA INSTITUTE OF ENGINEERING TECHNOLOGY,
GARIVIDI, ANDHRA PRADESH, INDIA

UPPE NANJI

ASSOC. PROFESSOR & HOD
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
ST.THERESSA INSTITUTE OF ENGINEERING TECHNOLOGY,
GARIVIDI, ANDHRA PRADESH, INDIA

Abstract—In modern electronic distribution networks, message authentication is an important objective of information security. This objective is met by providing the receiver of a message an assurance of the sender's identity. As physical protection such as sealed envelopes is not possible for messages expressed as binary sequences, digital tools have been developed using cryptography. A major limitation of all cryptographic methods for message authentication lies in their use of algorithms with fixed symmetric or public keys. In this paper, we propose an authenticated secret key transfer scheme that KDC (Key Distribution Center) can broadcast group key information to all group members at once and only authorized group members can recover the group key. The confidentiality of this transformation is information theoretically secure.

Index Terms— ciphers, data integrity, digital signature, encryption, key transport, message authentication, public-key cryptography, prepositioned secret sharing.

I. INTRODUCTION

With the rapid development of Internet technology and the popularization of multicast, group-oriented applications, such as video conference, network games, and video on demand, etc., are playing important roles. How to protect the communication security of these applications are also becoming more and more significant. Generally speaking, a secure group communication system should not only provide data confidentiality, user authentication, and information integrity, but also own perfect scalability. It is shown that a secure, efficient, and robust group key management approach is essential to a secure group

communication system. So, the secure storage of the private keys of a cryptosystem is an important problem. The possession of a highly sensitive key by an individual may not be desirable as the key can easily be lost or as the individual may not be fully trusted. Giving copies of the key to more than one individual increases the risk of compromise. A solution to this problem is to give shares of the key to several individuals, forcing them to cooperate to find the secret key. This not only reduces the risk of losing the key but also makes compromising the key more difficult.

In threshold cryptography, secret sharing deals with this problem, namely, sharing a highly sensitive secret among a group of n users so that only when a sufficient number t of them come together can the secret be reconstructed. Well-known secret sharing schemes (SSS) in the literature include Shamir [1] based on polynomial interpolation, Blakley [2] based on hyper plane geometry, and Asmuth-Bloom [3] based on the Chinese Remainder Theorem.

A shortcoming of secret sharing schemes is the need to reveal the secret shares during the reconstruction phase. The system would be more secure if the subject function can be computed without revealing the secret shares or reconstructing the secret. This is known as the function sharing problem. A function sharing scheme requires distributing the function's computation according to the underlying SSS such that each part of the computation can be carried out by a different user and then the partial results can be combined to yield the function's value without disclosing the individual secrets. Several protocols for function sharing have been proposed in the

literature [4,5,6,7]. Nearly all the existing solutions for function sharing uses Shamir secret sharing as the underlying SSS.

II. RELATED WORK

Several good explorations have been done for dealing with the group key distribution in a large group with frequent membership changes. There are two types of key establishment protocols: key transfer protocols and key agreement protocols. Key transfer protocols rely on a mutually trusted key generation center (KGC) to select session keys and then transport session keys to all communication entities secretly. Most often, KGC encrypts session keys under another secret key shared with each entity during registration.

In key agreement protocols, all communication entities are involved to determine session keys. The common key agreement protocol used in most distributed group key management protocols is Diffie-Hellman (DH) key agreement protocol. Some of the examples are: Bresson et al. [8] constructed a generic authenticated group DH Key exchange and the algorithm is provably secure. Katz and Yung [9] proposed the first constant-round and fully scalable group DH protocol which is provably secure in the standard model. The main feature of the group DH key exchange is to establish a secret group key among all group members without relying on a mutually trusted KGC.

A. Secret Sharing Schemes

The problem of secret sharing and the first solutions were introduced in 1979 independently by Shamir [1] and Blakley [2]. A $(t; n)$ -secret sharing scheme is used to distribute a secret d among n people such that any coalition of size t or more can construct d but smaller coalitions cannot. Shamir secret sharing is based on polynomial interpolation over a finite field. It uses the fact that we can find a polynomial of degree $t-1$ given t data points. Blakley secret sharing scheme has a different approach based on hyperplane geometry: To implement a $(t; n)$ threshold scheme, each of the n users is given a hyperplane equation in a t dimensional space over a finite field such that each hyperplane passes through a certain point. The intersection point of the hyperplanes is the secret. When t users come together, they can solve the system of equations to find the secret.

B. Function Sharing Schemes

Function sharing is the concept of distribution of the computation of a function such that when a sufficient number of users come together they can compute the value of the function without revealing their secret shares but less than the threshold number of users cannot. This problem is related to

secret sharing as the secret values needed for partial computations are distributed using secret sharing. Several solutions for sharing the RSA, ElGamal, and Paillier private key operations have been proposed in the literature [33,4,5,6,7]. Almost all of these schemes have been based on the Shamir SSS. The additive nature of the Lagrange's interpolation formula used in the combining phase of Shamir's scheme makes it an attractive choice for function sharing, but it also provides several challenges. One of the most significant challenges is the computation of inverses for the division operations in Lagrange's formula

III. SYSTEM ARCHITECTURE

In this section, we first describe the model of our proposed secret transfer protocol. Then we present the security goals of our group transfer protocol.

Each individuals of the group gets the key from the distribution center which will be exchanged between other members of the group.

- Key generation Each member of a group sends a request to KDC with a nonce.
- KDC collects a composite number factorizable in to prime numbers of the form $(4n+3)$. These numbers are obtained from the expression of Vandermonde's determinant.
- Random prime number is selected and being sent to each member of the group
- This number is represented as cyclic code with parity check and checksum.
- A cyclic shift is made on each representation.

In Fig.1, KDC is the key distribution center and has three members A1, A2 and A3. Group member A1 sends request IDA/N2 to KDC, KDC grants key to A1 by a reply IDA/N1. Similarly all other group members getting keys from KDC and communicate with KDC and with other group members. Life time of the key depends upon the number of transactions or sessions of the transaction.

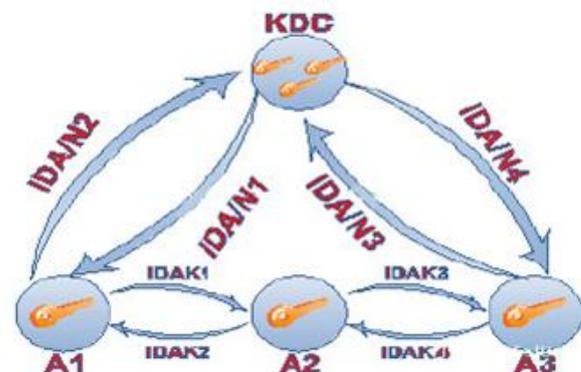


Fig 1 Key distribution

A. Goals

The main security goals of our group key transfer protocols are key authentication, key freshness and key authentication.

Each group key has never been used earlier in any step of KDC it ensures key freshness and cannot cause any further damage of group communication. Authorized group member only can recover the group key, it ensures that key confidentiality. KDC is giving assurance that the authorized group key is distributed to the group members, not by an intruder.

From our security analysis, we prove that none of the inside and outside attacks can successfully attack the authorized group members since attackers can neither obtain the group key nor share a group key with authorized group members.

B. Security Analysis

In this section, we prove that our proposed protocol achieves the security goals and is against inside as well as outside attacks. The two types of adversaries are outsiders and insiders. The outsider can try to recover the group key by impersonating as a group member by giving request to the KDC. In security analysis we will show that the outside attacker gains nothing, because the attacker cannot recover the group key, because they could not gain the individual factors of the composite number used by the KDC and the prime number difference are alone known from the Vandermonde's determinant

evaluation which are the public information available to the outsiders. The individual keys are generated under cyclic permutation and cyclic code representation, getting information may not help decoding permutation and cyclic code radix. Hence the inside attacker

IV. OUR CONTRIBUTION

Although a contributory group key agreement is a promising solution to achieve access control in collaborative and dynamic group applications, the existing schemes have not achieved the performance lower bound in terms of time, communication, and computation costs. In tree-based contributory group key agreement schemes, keys are organized in a logical tree structure, referred to as the key tree. In a key tree, the root node represents the group key, leaf nodes represent the members' private keys, and each intermediate node corresponds to a subgroup key shared by all the members (leaf nodes) under this node. The key of each non-leaf node is generated by performing the two-party DH between the two subgroups represented by its two children where each child represents the subgroup including all the members (leaf nodes) under this node.

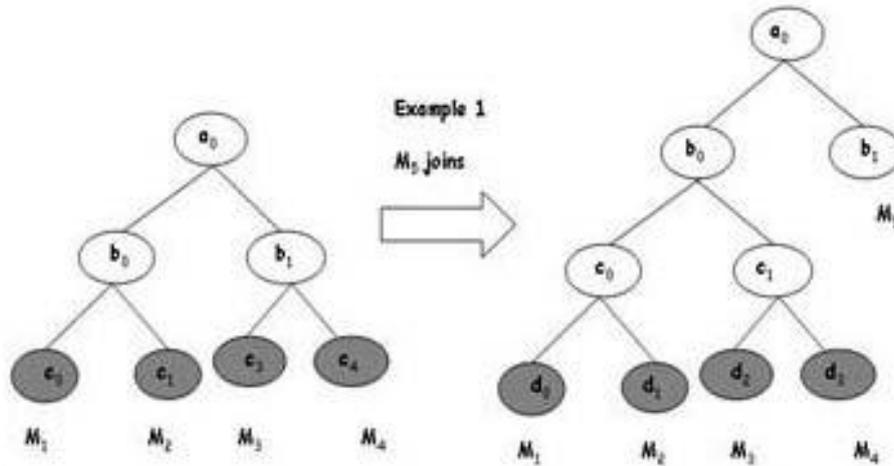


Fig.2 An example1 of key tree update upon single-user join event

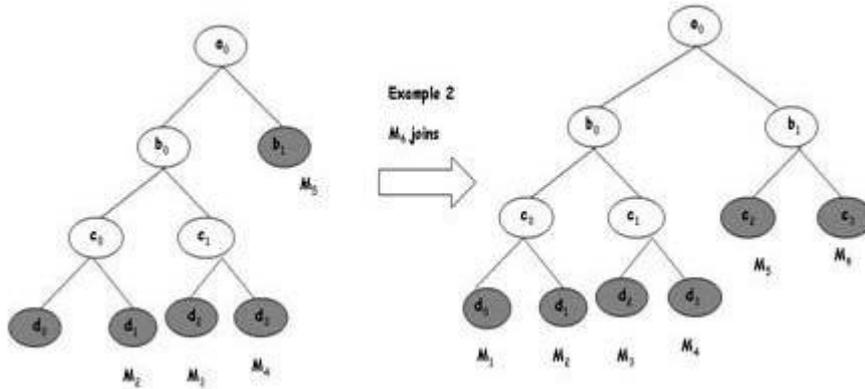


Fig.3 An example2 of key tree update upon single-user join event

A. Rekeying on Single-User Join

When a new user M wants to join the group G, the PACK initiates the single-user join protocol by broadcasting a request message that contains its member ID, a join request, its own blinded key, some necessary authentication information, and its signature for this request message. After receiving this user join request message the current group members will check and a new group key will be generated in order to incorporate a secret share from M. The rekeying upon single-user joins needs to perform two rounds of MDS code. Fig.2,3 shows two examples of a key tree update upon single-user join events.

In the first example tree consists of four members. After the new member M5 joins the group, a new node is created to act as the new root, and the node (b1) becomes the new join tree that represents M5. In the second example, when M6 joins the group, at the first round, the MDS is first performed between M5 and M6 to generate a new join tree, at the

second round, the MDS is performed between the new join tree and the main tree to generate a new group key.

B. Rekeying on Single-User Leave

When a current group member Y wants to leave the group, it broadcasts a leave request message to initiate the single user leave protocol, which contains its ID, a leave request, and a signature for this message. In order to reduce the rekeying cost upon a single-user leave event, PACK creates a phantom node that allows an existing member to simultaneously occupy more than one leaf node in the key tree. Fig 4 depicts the model of user leave and in this example, user M6 leaves the group where node (b0) is the root of the main tree and node (b1) is the root of the join tree. Since the size of the join tree is 2, the node representing M6 will be directly removed from the key tree, M5 changes its secret share, and a new group key will be generated by applying the MDS between M5 and the subgroup in the main tree.

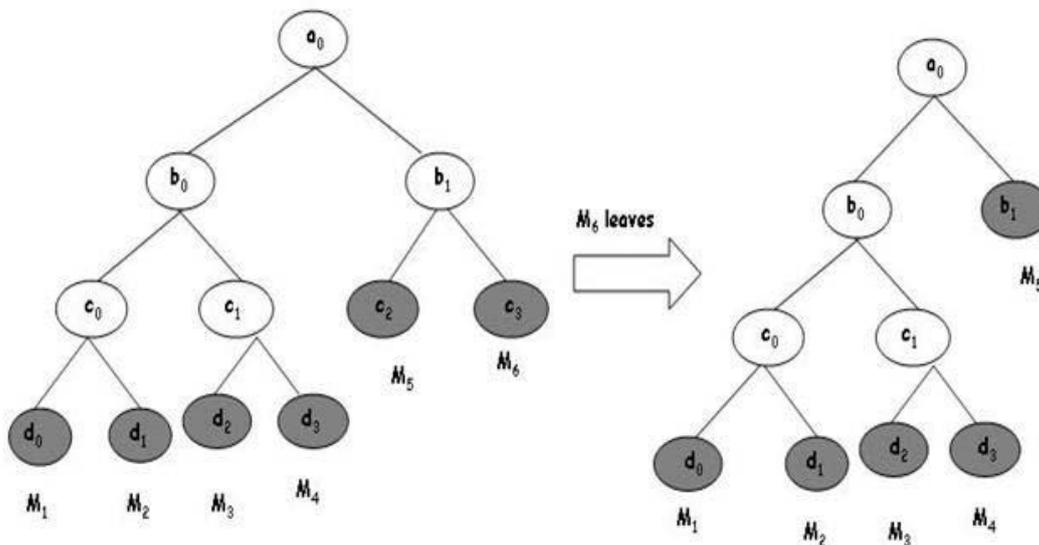


Fig.4 An example of key tree update upon single-user leave event

A. Rekeying on Multi-user Join and Leave protocol

PACK also has group merge and group partition protocols to handle simultaneously the join and leave of multiple users. Although multiple user events can be implemented by applying a sequence of single-user join or leave protocols, such sequential implementations are usually not cost-efficient. The group merges protocol, combines two or more groups into a single group, and returns a PF key tree. Group partition protocol, removes multiple group members simultaneously from the current group and construct a new PF key tree for the rest of the group members. In the group merge protocols, after removing all phantom nodes from those key trees corresponding to different subgroups, each key tree is split into several full key trees. The final result is obtained by uniting these full key trees into a PF tree using unite procedure. Similar to the group partition protocol, after removing all phantom nodes and leaving nodes, the original key tree is split into several full key trees, and the unite procedure is then applied on these full key trees to create a PF key tree. Since the height of the returned tree is $\log n$, where n is the group size after merging/partitioning, the time cost of group merge/partition is bounded by $O(\log n)$.

V. RESULTS

The experiments are carried out on an Intel Core 2 Duo 2.80-GHz machine with a 2-Gbyte memory running windows XP. The implementation results of computations and communications are presented, from these results; we can see that upon a single-user join event, proposed method has the lowest cost among all the schemes. Compared with other methods, proposed method has more than 10 percent reduction in computation cost and a more than 65 percent reduction in communication cost and time cost. Compared with GC, the reduction is even more, about 50 percent in computation cost and about 80 percent in time and communication costs. Upon a single-user leave event, compared with other methods, proposed method has about a 25 percent reduction in computation cost, about a 15 percent reduction in time cost, and a similar communication cost. Although other method has slightly higher computation and communication costs than other methods upon a single user leave event, when averaged over both join and leave events, the reduction is still significant, with a 20 percent reduction in computation cost, 35 percent reduction in communication cost, and 40 percent reduction in time cost. Fig 5 and 6, shows the key distribution time and key recovery time of both the scheme under various multicast group sizes. It is clear that using one-way hash functions adds none-trivial computation complexity.

VI. CONCLUSIONS AND FUTURE WORK

For any multicast group communication, group key agreement was found to be challenging because of its dynamic nature. Group key management scheme are of either distributed, centralized or hybrid architecture. Although many solutions have been proposed to handle group key changes, In this paper, We have proposed an efficient group key transfer protocol based on secret sharing. Every user

needs to register at a trusted KGC initially and preshare a secret with KGC. KGC broadcasts group key information to all group members at once. The confidentiality of our group key distribution is information theoretically secure. This scheme is thus practical for many applications in various broadcast capable networks such as Internet and wired and wireless networks.

Future Work

As mentioned above related work still there is a necessity to provide the security in the network, so there is a future scope to enhancements.

REFERENCES

- [1] Shamir. How to share a secret? *Comm. ACM*, 22(11):612–613, 1979.
- [2] C. Asmuth and J. Bloom. A modular approach to key safeguarding. *IEEE Trans. Information Theory*, 29(2):208–210, 1983.
- [3] G. Blakley. Safeguarding cryptographic keys. In *Proc. of AFIPS National Computer Conference*, 1979.
- [4] Y. Desmedt. Some recent research aspects of threshold cryptography. In *Proc. of ISW '97, 1st International Information Security Workshop*, volume 1196 of *LNCS*, pages 158–173. Springer-Verlag, 1997.
- [5] Y. Desmedt and Y. Frankel. Threshold cryptosystems. In *Proc. of CRYPTO '89*, volume 435 of *LNCS*, pages 307–315. Springer-Verlag, 1990.
- [6] Y. Desmedt and Y. Frankel. Shared generation of authenticators and signatures. In *Proc. of CRYPTO '91*, volume 576 of *LNCS*, pages 457–469. Springer-Verlag, 1992.
- [7] Y. Desmedt and Y. Frankel. Homomorphic zero-knowledge threshold schemes over any finite abelian group. *SIAM Journal on Discrete Mathematics*, 7(4):667–679, 1994.
- [8] P. A. Fouque, G. Poupard, and J. Stern. Sharing decryption in the context of voting or lotteries. In *Proc. of FC 2000, 4th International Conference on Financial Cryptography*, volume 1962 of *LNCS*, pages 90–104. Springer-Verlag, 2001.
- [9] E. Bresson, O. Chevassut, and D. Pointcheval. “Provably-Secure Authenticated Group Diffie-Hellman Key Exchange,” *ACM Trans. Information and System Security*, vol. 10, no. 3, pp. 255-264, Aug. 2007.
- [10] J. Katz and M. Yung, “Scalable Protocols for Authenticated Group Key Exchange,” *J. Cryptology*, vol. 20, pp. 85-113, 2007.

IMPLEMENTATION OF OBJECT ORIENTED APPROACH TO FAST IP RECOVERY BASED ON MULTIPLE ROUTING CONFIGURATIONS

HARSHA PUJARI

2/2 M.TECH CSE, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

U.D.PRASANNA

ASSOC.PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

Abstract— To assure fast recovery from link and node failures in IP networks, we present JAVA based recovery scheme called Multiple Routing Configurations (MRC). This proposed scheme guarantees recovery in all single failure scenarios, using a single mechanism to handle both link and node failures, and without knowing the root cause of the failure. MRC is based on keeping additional routing information in the routers, and allows packet forwarding to continue on an alternative output link immediately after the detection of a failure.

Index Terms— MRC; Java; Node,

I. INTRODUCTION

In recent years the Internet has been transformed from a special purpose network to a ubiquitous platform for a wide range of everyday communication services. The demands on Internet reliability and availability have increased accordingly. A disruption of a link in central parts of a network has the potential to affect hundreds of thousands of phone conversations or TCP connections, with obvious adverse effects. The ability to recover from failures has always been a central design goal in the Internet. This re-convergence assumes full distribution of the new link state to all routers in the network domain. When the new state information is distributed, each router individually calculates new valid routing tables.

This network-wide IP re-convergence is a time consuming process, and a link or node failure is typically followed by a period of routing instability. During this period, packets may be dropped due to invalid routes. This phenomenon has been studied in IGP [1] and has an adverse effect on real-time applications [2]. Events leading to a re-convergence have been shown to occur frequently [3]. Much effort has been devoted to optimizing the different

steps of the convergence of IP routing, i.e., detection, dissemination of information and shortest path calculation, but the convergence time is still too large for applications with real time demands. A key problem is that since most network failures are short lived, too rapid triggering of the re-convergence process can cause route flapping and increased network instability.

The IGP convergence process is slow because it is reactive and global. It reacts to a failure after it has happened, and it involves all the routers in the domain. In this project I present a new scheme for handling link and node failures in IP networks. Multiple Routing Configurations (MRC) is proactive and local, which allows recovery in the range of milliseconds. MRC allows packet forwarding to continue over pre-configured alternative next-hops immediately after the detection of the failure. Using MRC as a first line of defense against network failures, the normal IP convergence process can be put on hold. This process is then initiated only as a consequence of non-transient failures. MRC guarantees recovery from any single link or node failure, which constitutes a large majority of the failures experienced in a network [5]. MRC makes no assumptions with respect to the root cause of failure, e.g., whether the packet forwarding is disrupted due to a failed link or a failed router.

The main idea of MRC is to use the network graph and the associated link weights to produce a small set of backup network configurations. The link weights in these backup configurations are manipulated so that for each link and node failure, and regardless of whether it is a link or node failure, the node that detects the failure can safely forward the incoming packets towards the destination on an alternate link. MRC assumes that the network uses shortest path routing and destination based hop-by-hop forwarding.

The shifting of traffic to links bypassing the failure can lead to congestion and packet loss in parts of the network [6]. This limits the time that the

proactive recovery scheme can be used to forward traffic before the global routing protocol is informed about the failure, and hence reduces the chance that a transient failure can be handled without a full global routing re-convergence. Ideally, a proactive recovery scheme should not only guarantee connectivity after a failure, but also do so in a manner that does not cause an unacceptable load distribution. This requirement has been noted as being one of the principal challenges for pre-calculated IP recovery schemes [7]. With MRC, the link weights are set individually in each backup configuration. This gives great flexibility with respect to how the recovered traffic is routed. The backup configuration used after a failure is selected based on the failure instance, and thus we can choose link weights in the backup configurations that are well suited for only a subset of failure instances.

It is important to stress that MRC does not affect the failure free original routing, i.e., when there is no failure, all packets are forwarded according to the original configuration, where all link weights are normal. Upon detection of a failure, only traffic reaching the failure will switch configuration. All other traffic is forwarded according to the original configuration as normal. If a failure lasts for more than a specified time interval, a normal re-convergence will be triggered. MRC does not interfere with this convergence process, or make it longer than normal. However, MRC gives continuous packet forwarding during the convergence, and hence makes it easier to use mechanisms that prevent micro-loops during convergence, at the cost of longer convergence times. If a failure is deemed permanent, new configurations must be generated based on the altered topology.

II. RELATED WORK

There are several proposals for mitigating the impact of link failures on network performance. MRC is to use the network graph and the associated link weights to produce a small set of backup network configurations. The link weights in these backup configurations are manipulated so that for each link and node failure, and regardless of whether it is a link or node failure, the node that detects the failure can safely forward the incoming packets towards the destination. MRC assumes that the network uses shortest path routing and destination based hop-by-hop forwarding.

In this thesis, it is sometimes claimed that the node failure recovery implicitly addresses link failures too, as the adjacent links of the failed node can be avoided. This is true for intermediate nodes, but the destination node in a network path must be

reachable if operative ("The last hop problem", [6]). MRC solves the last hop problem by strategic assignment of link weights between the backup configurations.

MRC has a range of attractive features:

- It gives almost continuous forwarding of packets in the case of a failure. The router that detects the failure initiates a local rerouting immediately, without communicating with the surrounding neighbors.
- MRC helps improve network availability through suppression of the re-convergence process. Delaying this process is useful to address transient failures, and pays off under many scenarios [4]. Suppression of the re-convergence process is further actualized by the evidence that a large proportion of network failures is short-lived, often lasting less than a minute [5].
- MRC uses a single mechanism to handle both link and node failures. Failures are handled locally by the detecting node, and MRC always finds a route to the destination (if operational).
- MRC makes no assumptions with respect to the root cause of failure, e.g., whether the packet forwarding is disrupted due to a failed link or a failed router. Regardless of this, MRC guarantees that there exists a valid, preconfigured next-hop to the destination.
- An MRC implementation can be made without major modifications to existing IGP routing standards. IETF recently initiated specifications of multi-topology routing for OSPF and IS-IS, and this approach seems well suited to implement our proposed backup configurations [7][8][9]. The concept of multiple routing configurations and its application to network recovery is not new. Our main inspiration has been a layer-based approach used to obtain deadlock-free and fault-tolerant routing in irregular cluster networks based on a routing strategy called Up*/Down*[10].

General packet networks are not hampered by deadlock considerations necessary in interconnection networks, and hence I generalized the concept in a technology independent manner and named it Resilient Routing Layers [11]. In the graph-theoretical context, RRL is based on calculating spanning sub topologies of the network, called layers. Each layer contains all nodes but only a subset of the links in the network.. The work described in this paper differs substantially from RRL in that, we do not alter topologies by removing links, but rather manipulate link weights to meet goals of handling both node and link failures without needing to know the root cause of the failure. In MRC, all links remain in the topology,

but in some configurations, some links will not be selected by shortest path routing mechanisms due to high weights.

III. Theoretical Background

Network topology is the study of the arrangement or mapping of the elements (links, nodes, etc.) of a network, especially the physical (real) and logical (virtual) interconnections between nodes.

Much work has lately been done to improve robustness against component failures in IP networks [7]. In this section, I focus on the most important contributions aimed at restoring connectivity without a global re-convergence. This indicates whether each mechanism guarantees one-fault tolerance in an arbitrary bi-connected network, for link and node failures, independent of the root cause of failure (failure agnostic). This also indicates whether MRC solve the “last hop problem”.

Network layer recovery in the timescale of milliseconds has traditionally only been available for networks using MPLS with its fast reroute extensions [12]. In the discussion below, I focus mainly on solutions for connectionless destination-based IP routing. IETF has recently drafted a framework called IP fast reroute where they point at Loop-Free Alternates (LFAs) as a technique to partly solve IP fast reroute. From a node detecting a failure, a next hop is defined as an LFA if this next hop will not loop the packets back to the detecting node or to the failure. Since LFAs do not provide full coverage, IETF is also drafting a tunneling approach based on so called “Not-via” addresses to guarantee recovery from all single link and node failures [8]. Not-via is the connectionless version of MPLS fast reroute [12] where packets are detoured around the failure to the next-next hop.

To protect against the failure of a component P, a special not-via address is created for this component at each of P's neighbors. Forwarding tables are then calculated for these addresses without using the protected component. This way, all nodes get a path to each of P's neighbors, without passing through (“Not-via”) P. The Not-via approach is similar to MRC in that loop-free backup next-hops are found by doing shortest path calculations on a subset of the network. It also covers against link and node failures using the same mechanism and is strictly pre-configured. However, the tunneling approach may give less optimal backup paths, and less flexibility with regards to post failure load balancing. Narvaez et al. [13] propose a method relying on multi-hop repair paths. They propose to do a local re-convergence upon

detection of a failure, i.e., notify and send updates only to the nodes necessary to avoid loops.

A similar approach also considering dynamic traffic engineering is presented in [14]. I call these approaches local rerouting. They are designed only for link failures, and therefore avoid the problems of root cause of failure and the last hop. Their method does not guarantee one-fault-tolerance in arbitrary bi-connected networks. It is obviously connectionless. However, it is not strictly pre-configured, and can hence not recover traffic in the same short time-scale as a strictly pre-configured scheme. Nelakuditi et al. [4] propose using interface specific forwarding to provide loop-free backup next hops to recover from link failures.

Their approach is called failure insensitive routing (FIR). The idea behind FIR is to let a router infer link failures based on the interface packets are coming from. When a link fails, the attached nodes locally reroute packets to the affected destinations, while all other nodes forward packets according to their pre-computed interface specific forwarding tables without being explicitly aware of the failure. In another paper, they have also proposed a similar method, named Failure Inference based Fast Rerouting (FIFR), for handling node failures [15]. This method will also cover link failures, and hence it operates independent of the root cause of failure. However, their method will not guarantee this for the last hop, i.e., they do not solve the “last hop problem”. FIFR guarantees one-fault-tolerance in any bi-connected network, it is connectionless, pre-configured and it does not affect the original failure-free routing. Our main inspiration for using multiple routing functions to achieve failure recovery has been a layer-based approach used to obtain deadlock-free and fault-tolerant routing in irregular cluster networks [16].

General packet networks are not hampered by deadlock considerations necessary in interconnection networks, and hence we generalized the concept in a technology independent manner and named it Resilient Routing Layers [17]. In the graph-theoretical context, RRL is based on calculating spanning sub topologies of the network, called layers. Each layer contains all nodes but only a subset of the links in the network. In this paper I refine these ideas and adapt them to an IP setting. None of the proactive recovery mechanisms discussed above takes any measures towards a good load distribution in the network in the period when traffic is routed on the recovery paths.

Existing work on load distribution in connectionless IGP networks has either focused on the failure free case [18] or on finding link weights that work well both in the normal case and when the

routing protocol has converged after a single link failure [19]. Many of the approaches listed provide elegant and efficient solutions to fast network recovery, however MRC and Not-via tunneling seems to be the only two covering all evaluated requirements. However, MRC offers the same functionality with a simpler and more intuitive approach, and leaves more room for optimization with respect to load balancing.

IV. SYSTEM ARCHITECTURE

To send the packets from source node to destination node, first it checks the neighbor nodes of source node. The source node requests the destination node to generate the available paths. If don't select the destination node it asks to give destination node. To select the particular shortest path, and send the packets to the particular destination node. The failures are fairly common in the everyday operation of a network due to various

causes such as maintenance, fault interfaces, and accidental fiber cuts.

MRC is based on using a small set of backup routing configurations, where each of them is resistant to failures of certain nodes and links. The original network topology, a configuration is defined as a set of associated link weights. In a configuration that is resistant to the failure of a particular node n , link weights are assigned so that traffic routed according to this configuration is never routed through node n . The failure of node n then only affects traffic that is sent from or destined to n . Similarly, in a configuration that is resistant to failure of a link l , traffic routed in this configuration is never routed over this link, hence no traffic routed in this configuration is lost if l fails. In MRC, node n and link l are called isolated in a configuration, when, as described above, no traffic routed according to this configuration is routed through n or l .

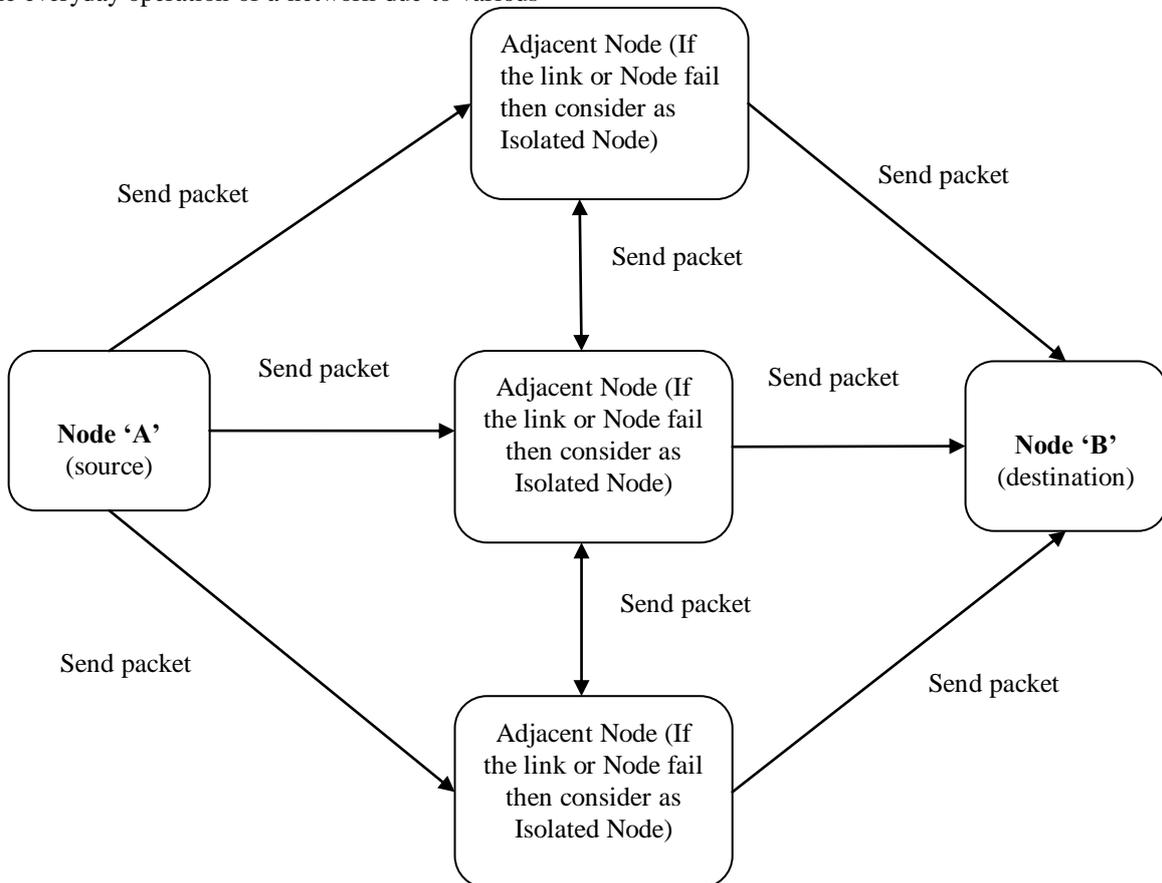


Fig. 1. System Architecture

MRC approach is threefold. First, create a set of backup configurations, so that every network component is isolated in one configuration. Second, for each configuration, a standard routing algorithm like OSPF is used to calculate configuration specific shortest path trees and create forwarding tables in

each router, based on the configurations. The use of a standard routing algorithm guarantees loop free forwarding within one configuration. Forwarding process that takes advantage of the backup configurations to provide fast recovery from a component failure.

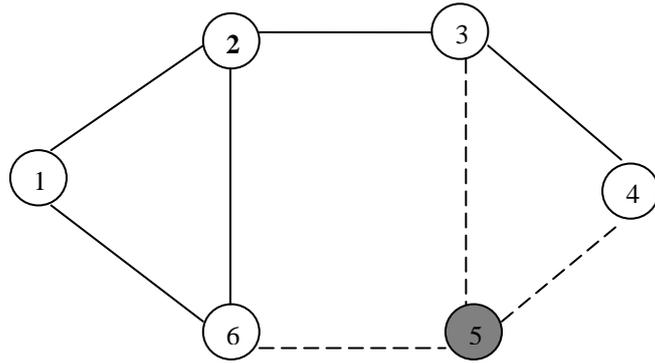


Fig. 1. Isolating a node.

Fig.2 illustrates a configuration where node 5 is isolated. In this configuration, the weight of the stapled links is set so high that only traffic sourced by or destined for node 5 will be routed over these links, which are restricted links. Node failures can be handled through blocking the node from transiting traffic. This node-blocking will normally also protect the attached links. But a link failure in

the last hop of a path can obviously not be recovered by blocking the downstream node (ref. “the last hop problem”). Hence, I must make sure that, in one of the backup configurations, there exists a valid path to the last hop node, without using the failed link. A link is isolated by setting the weight to infinity, so that any other path would be selected before one including that link.

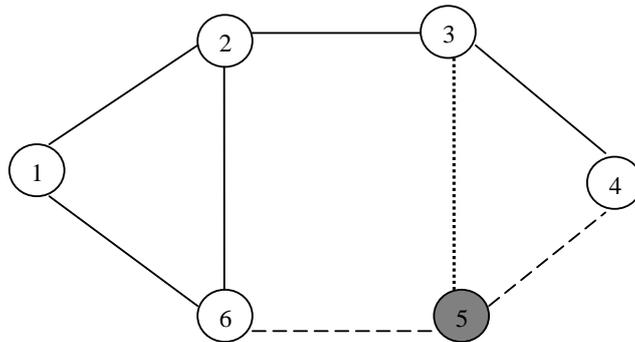


Fig. 3 Isolating links of a node.

Fig.3 shows the same configuration as before, except now link 3-5 has been isolated (dotted). No traffic is routed over the isolated link in this configuration; traffic to and from node 5 can only use the restricted links. In Fig.4, shows how several

nodes and links can be isolated in the same configuration. In a backup configuration like this, packets will never be routed over the isolated (dotted) links, and only in the first or the last hop be routed over the restricted (dashed) links.

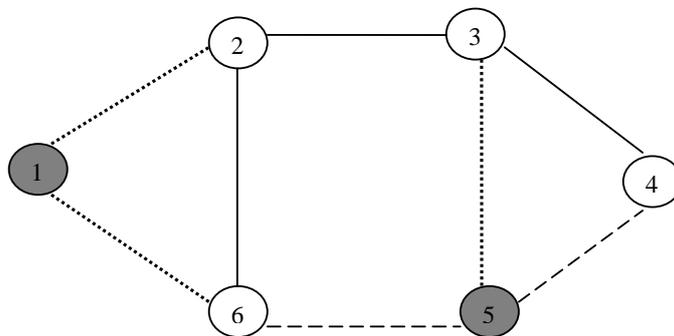


Fig.4 isolating more than one node

Some important properties of a backup configuration are worth pointing out. First, all non-isolated nodes are internally connected by a sub-graph that does not contain any isolated or restricted links. We denote this sub-graph as the

backbone of the configuration. In the backup configuration shown in Fig.4, nodes 6, 2 and 3 with their connecting links constitute this backbone. Second, all links attached to an isolated node are either isolated or restricted, but an isolated node is

always directly connected to the backbone with at least one restricted link. Using a standard shortest path calculation, each router creates a set of configuration-specific forwarding tables. For simplicity, we say that a packet is forwarded according to a configuration, meaning that it is forwarded using the forwarding table calculated based on that configuration.

When a router detects that a neighbor can no longer be reached through one of its interfaces, it does not immediately inform the rest of the network about the connectivity failure. Instead, packets that would normally be forwarded over the failed interface are marked as belonging to a backup configuration, and forwarded on an alternative interface towards its destination. The packets must be marked with a configuration identifier, so the routers along the path know which configuration to use. Packet marking is most easily done by using the DSCP field in the IP header.

If this is not possible, other packet marking strategies like IPv6 extension headers or using a private address space and tunneling can be imagined. It is important to stress that MRC does not affect the failure free original routing, i.e. when there is no failure, all packets are forwarded according to the original configuration, where all link weights are normal. Upon detection of a failure, only traffic reaching the failure will switch configuration. All other traffic is forwarded according to the original configuration as normal.

A. Generating Backup Configurations

The algorithm will typically be run once at the initial startup of the network, and each time a node or link is permanently added or removed to give on the back up configurations used in MRC.

Configuration Constraints

To guarantee single-failure tolerance and consistent routing, the backup configurations used in MRC must adhere to the following requirements:

1. A node must not carry any transit traffic in the configuration where it is isolated. Still, traffic must be able to depart from and reach an isolated node.
2. A link must not carry any traffic at all in the configuration where it is isolated.
3. In each configuration, all node pairs must be connected by a path that does not pass through an isolated node or an isolated link.

Every node and every link must be isolated in at least one backup configuration.

The first requirement decides what weights must be put on the restricted links attached to an isolated node. To guarantee that no path will go through an isolated node, it suffices that the restricted links

have a weight W of at least the sum of all link weights w in the original configuration

$$W > \sum_{e_{i,j} \in E} w_{i,j}$$

It guarantees that any other path between two nodes in the network will be chosen by a shortest path algorithm before one passing through the isolated node. Only packets sourced by or destined for the isolated node itself will traverse a restricted link with weight W , as they have no shorter path. An algorithm, restricted and isolated links are given the same weight in both directions in the backup configurations, i.e., we treat them as undirected links. However, it does not prevent the use of independent link weights in each direction in the default configuration. The second requirement implies that the weight of an isolated link must be set so that traffic will never be routed over it. Such links are given infinite weight.

V. IMPLEMENTATION

In this paper, we consider four modules.

- Network construction
- Find Link Failure
- Calculate Load Balancing
- Find Isolated Node

A. Network construction

MRC configurations are defined by the network topology, which is the same in all configurations, and the associated link weights, which differ among configurations. We formally represent the network topology as a graph, with a set of nodes and a set of unidirectional links. In order to guarantee single-fault tolerance, the topology graph must be bi-connected. A configuration is defined by this topology graph and the associated link weight function.

B. Find Link Failure

Send Packets through constructed network towards Destination Node. If sending node receive acknowledgement from Destination, means the link that forward packet is good. If not the link is considered as failure. This failure link is also called as isolated link.

C. Calculate Load Balancing

MRC offers functionality with a simpler and more intuitive approach, and leaves more room for optimization with respect to load balancing.

The backup configurations are constructed in a way that gives better load balancing and avoids congestion after a failure. We propose a procedure to do this by constructing a complete set of valid configurations in three phases. First, the link

weights in the normal configuration are optimized for the given demand matrix while only taking the failure free situation into account. Second, we take advantage of the load distribution in the failure free case to construct the MRC backup configurations in an intelligent manner. Finally, we optimize the link weights in the backbones of the backup configurations to get a good load distribution after any link failure.

D. Find Isolated Node

A node must not carry any transit traffic in the configuration where it is isolated. Still, traffic must be able to depart from and reach an isolated node.

With MRC, restricted links are always attached to isolated nodes.

A restricted link connects the node before failure link and find alternate path, by searching this isolated node. Isolated node is the node which does not carry any traffic. With our algorithm, all nodes and links in the network are isolated in exactly one configuration.

The third property above results in the following two invariants for our algorithm, which must be evaluated each time a new node and its connected links are isolated in a configuration.

- 1) A configuration must contain a backbone
- 2) All isolated nodes in a configuration must be directly connected to the backbone through at least one restricted link.

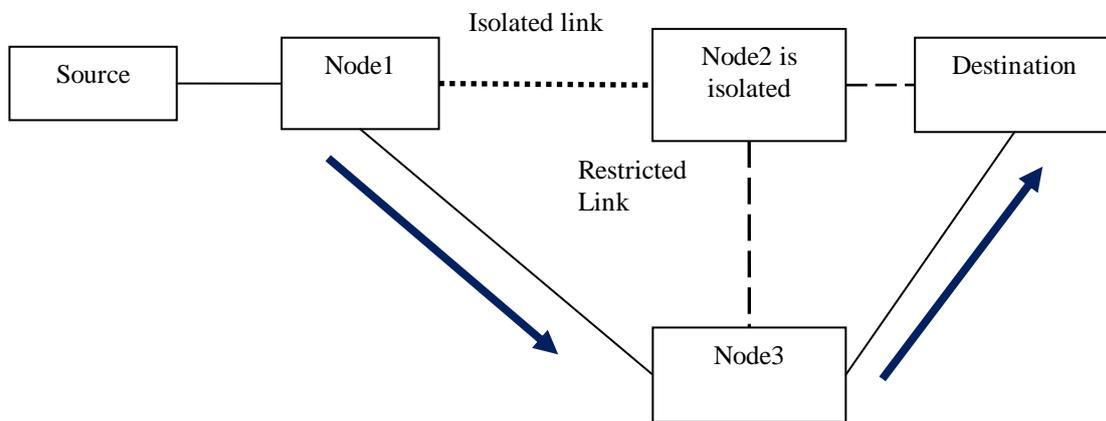


Fig.5 Finding isolate node

E. Performance Requirements

Performance is measured in terms of the output provided by the application. Requirement specification plays an important part in the analysis of a system. Only when the requirement specifications are properly given, it is possible to design a system, which will fit into required environment. It rests largely with the users of the existing system to give the requirement specifications because they are the people who finally use the system. This is because the requirements have to be known during the initial stages so that the system can be designed according to those requirements. It is very difficult to change the system once it has been designed and on the

other hand designing a system, which does not cater to the requirements of the user, is of no use.

VI. RESULTS

MRC requires the routers to store additional routing configurations. The amount of state required in the routers is related to the number of such backup configurations. Since routing in a backup configuration is restricted, MRC will potentially give backup paths that are longer than the optimal paths. Longer backup paths will affect the total network load and also the end-to-end delay. It must be noted that MRC yields the shown performance immediately after a failure. The complexity of the proposed algorithm is determined by worst case $O(|N|+|A|)$. Consider the following graph Fig.6

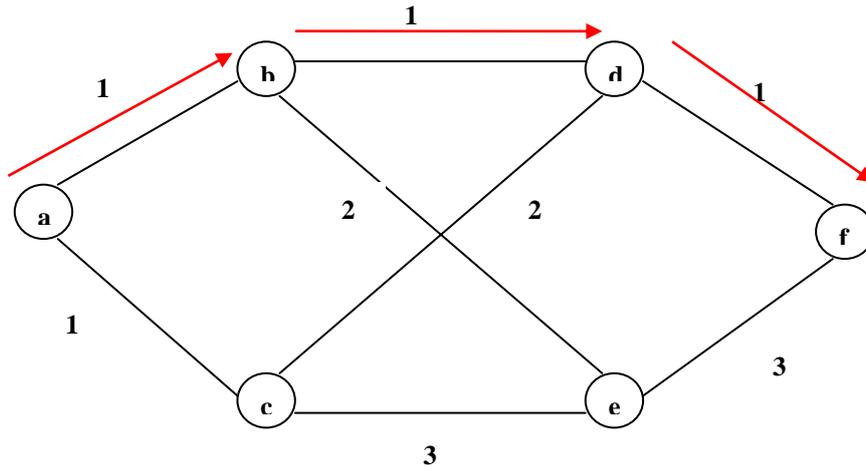


Fig.6.Configuration1

If data is send from source 'a' to destination 'f'. it choose the optimal path(a→b→d→f).Now I assume that the node 'd' is damaged, so isolate the node 'd' by isolating almost all its links and restricting at least one link, among the links to its

immediate neighbor present in the alternate path to destination 'f' , After isolating the node 'd', the backup configuration obtained 'C₁' is shown below Fig.7,.

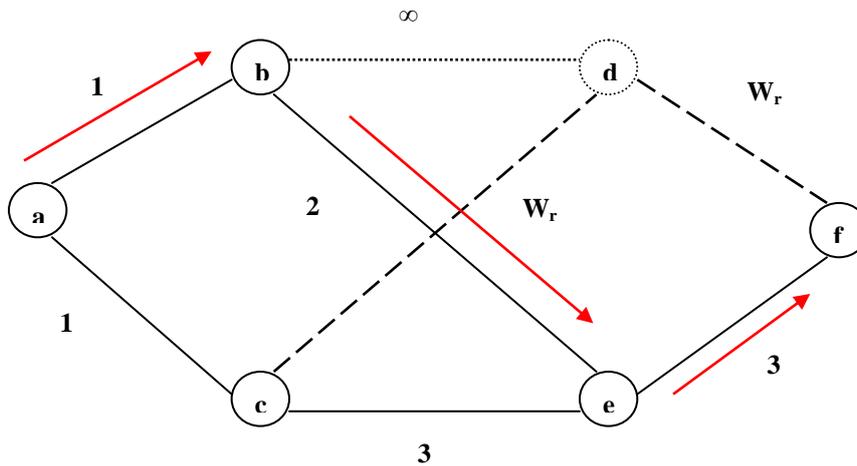


Fig.8.Configuration2

So the given complexity $O(|N|+|A|)$ has been proved to be true when a single node isolation is considered (d) which as per the above graph is $O(|1|+|3|)$. Now the data is send from source node 'a' to node 'f'. it

choose alternate path a→b→e→f. After isolating the node 'b', the backup configuration obtained 'C₂' shown below Fig.9, Now the alternate path a→c→e→f.

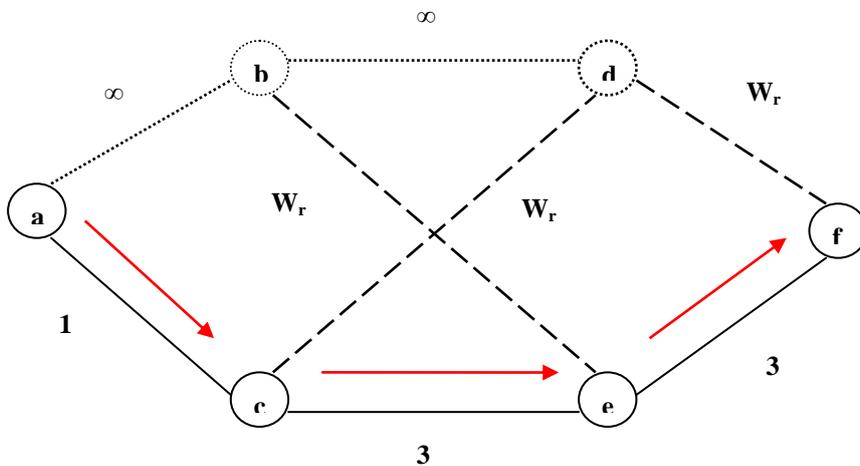


Fig.9.Configuration3

The complexity is also proved in the case of $N=2$ (nodes isolated are 'd' and 'b'). Which as per the above graph is $O(2+|5|)$. So the given computational complexity $O(N+|A|)$ can be verified easily, when we consider all the nodes in the graph for isolation.

VII. CONCLUSIONS AND FUTURE WORK

Multiple Routing Configurations as an approach to achieve fast recovery in IP networks. MRC guarantees recovery from any single node or link failure in an arbitrary bi-connected network. By calculating backup configurations in advance, and operating based on locally available information only, MRC can act promptly after failure discovery. MRC operates without knowing the root cause of failure, i.e., whether the forwarding disruption is caused by a node or link failure. This is achieved by using careful link weight assignment according to the rules we have described. The link weight assignment rules also provide basis for specification of a forwarding procedure.

In this project, I focused how the network can be used to improve the distribution of the recovered traffic, and thus reduce the chances of congestion when MRC is used.

Future Work:

From the viewpoint of networking, there are still lots of open problems. Among them, the following questions will be studied in future:

- To reduce the risk of congestion after a failure by doing traffic engineering through intelligent link weight assignment in each configuration.
- To maintaining a separate multicast tree for each configuration to achieve very fast recovery from both link and node failures.

REFERENCES

- [1] A. Basu and J. G. Riecke, "Stability Issues in OSPF Routing," in Proceedings of SIGCOMM 2001, pp. 225–236, August 2001
- [2] C. Boutremans, G. Iannaccone, and C. Diot, "Impact of link failures on VoIP performance," in Proceedings of International Workshop on Network and Operating System Support for Digital Audio and Video, 2002.
- [3] D. Watson, F. Jahanian, and C. Labovitz, "Experiences with monitoring OSPF on a regional service provider network," in ICDCS '03: Proceedings of the 23rd International Conference on Distributed Computing Systems. IEEE Computer Society, pp. 204–213, 2003.
- [4] S. Lee, Y. Yu, S. Nelakuditi, Z.-L. Zhang, and C.-N. Chuah, "Proactive vs. reactive approaches to failure resilient routing," in *Proceedings IEEE INFOCOM'04*, Mar. 2004.
- [5] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, and C. Diot, "Characterization of failures in an IP backbone network," in Proceedings of INFOCOM 2004, Mar. 2004.
- [6] S. Iyer, S. Bhattacharyya, N. Taft, and C. Diot, "An approach to alleviate link overload as observed on an IP backbone," in Proceedings of INFOCOM'03, pp. 406–416, Mar. 2003.
- [7] P. Psenak, S. Mirtorabi, A. Roy, L. Nguen, and P. Pillay-Esnault, "MTOSPF: Multi topology (MT) routing in OSPF," IETF Internet Draft, Apr. 2005.
- [8] T. Przygienda, N. Shen, and N. Sheth, "M-ISIS: Multi topology (MT) routing in IS-IS," Internet Draft, May 2005.
- [9] M. Menth and R. Martin, "Network resilience through multi-topology routing," University of Wurzburg, Institute of Computer Science, Tech. Rep. 335, May 2004.
- [10] I. Theiss and O. Lysne, "FROOTS - fault handling in up*/down* routed networks with multiple roots," in Proceedings of the International Conference on High Performance Computing, 2003
- [11] A. Kvalbein, A. F. Hansen, T. Cicic, S. Gjessing, and O. Lysne, "Fast recovery from link failures using resilient routing layers," in Proceeding 10th IEEE Symposium on Computers and Communications (ISCC), June 2005.
- [12] P. Pan, G. Swallow, and A. Atlas, "Fast reroute extensions to RSVP-TE for LSP tunnels," RFC 4090, May 2005.
- [13] P. Narvaez, K.-Y. Siu, and H.-Y. Tzeng, "Local restoration algorithms for link-state routing protocols," in Proc. IEEE Int. Conf. Computer Communications and Networks (ICCCN'99), pp. 352–357, Oct. 1999.
- [14] R. Rabbat and K.-Y. Siu, "Restoration methods for traffic engineered networks for loop-free routing guarantees," in Proc. IEEE Int. Conf. Communications (ICC'01), Helsinki, Finland, vol. 5, pp. 1566–1570, Jun. 2001.
- [15] Z. Zhong, S. Nelakuditi, Y. Yu, S. Lee, J. Wang, and C.-N. Chuah, "Failure inferencing based fast rerouting for handling transient link and node failures," in Proc. IEEE INFOCOM, vol. 4, pp. 2859–2863, Mar. 2005.
- [16] I. Theiss and O. Lysne, "FRoots, a fault tolerant and topology agnostic routing technique," IEEE Trans. Parallel Distrib. Syst., vol. 17, pp. 1136–1150, Oct. 2006.
- [17] A. F. Hansen, T. Cicic, S. Gjessing, A. Kvalbein, and O. Lysne, "Resilient routing layers for recovery in packet networks," in Proc. Int. Conf. Dependable Systems and Networks (DSN 2005), pp. 238–247, Jun. 2005.
- [18] B. Fortz and M. Thorup, "Internet traffic engineering by optimizing OSPF weights," in Proc. IEEE INFOCOM, pp. 519–528, 2000.
- [19] A. Sridharan and R. Guerin, "Making IGP routing robust to link failures," in Proc. Networking, Waterloo, Canada, 2005.

Implementation of object oriented approach to Index Support for Item Set Mining (IMine)

R.SRIKANTH,

2/2 M.TECH CSE, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI,
ANDHRA PRADESH, INDIA.

D.T.V.DHARMAJEE RAO

PROFESSOR & HOD,

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI,
ANDHRA PRADESH, INDIA.

Abstract— The increase in huge amount of data is seen clearly in present days because of requirement for storing more information. To extract certain data from this large database is a very difficult task. This leads to the researchers to drag themselves for developing better technique to mine the required data. There are various techniques proposed by several researchers to deal with this difficulty. Among various available techniques, association rule mining for extract the required data from the database is found to be better. This paper presents the IMine index, a general and compact structure which provides tight integration of item set extraction in a relational DBMS. Since no constraint is enforced during the index creation phase, IMine provides a complete representation of the original database. To reduce the I/O cost, data accessed together during the same extraction phase are clustered on the same disk block. Experiments, run for both sparse and dense data distributions, show the efficiency of the proposed index and its linear scalability also for large data sets.

Index Terms— Object Oriented Approach; IMine; Index Support; Item Set Mining.

I. INTRODUCTION

DATA mining is provoked by decision support difficulties featured by majority of business organizations and is illustrated as an significant field of research. One of the major difficulties in data mining is creating fast and efficient techniques that can deals with large volumes of data as majority mining techniques carry out computation over the complete database and frequently the databases are in huge size. Physical analysis of these huge amount of information stored in modern databases is very difficult. A recognized data mining technique is association rule mining. It is able to discover all interesting relationships which are called as associations in a database. Association rules are very efficient in revealing all the interesting relationships in a relatively large

database with huge amount of data. The large quantity of information collected through the set of association rules can be used not only for illustrating the relation-ships in the database, but also used for differentiating between different kinds of classes in a database. But the major difficulty in association rule mining is its complexity.

Research activity usually focuses on defining efficient algorithms for item set extraction, which represents the most computationally intensive knowledge extraction task in association rule mining [1]. The data to be analyzed is usually stored into binary files, possibly extracted from a DBMS. Most algorithms [2, 3] exploit ad hoc main memory data structures to efficiently extract item sets from a flat file. Recently, disk-based extraction algorithms have been proposed to support the extraction from large data sets [4,5,6], but still dealing with data stored in flat files. To reduce the computational cost of item set extraction, different constraints maybe enforced [7,8,9], among which the most simple is the support constraint, which enforces a threshold on the minimum support of the extracted item sets.

Relational DBMSs exploit indices, which are ad hoc data structures, to enhance query performance and support the execution of complex queries. In this paper, we propose a similar approach to support data mining queries. The IMine index (Item set-Mine index) is a novel data structure that provides a compact and complete representation of transactional data supporting efficient item set extraction from a relational DBMS. It is characterized by the following properties:

1. It is a covering index. No constraint (e.g., support constraint) is enforced during the index creation phase. Hence, the extraction can be performed by means of the index alone, without

accessing the original database. The data representation is complete and allows reusing the index for mining item sets with any support threshold.

2. The IMine index is a general structure which can be efficiently exploited by various item set extraction algorithms. These algorithms can be characterized by different in-memory data representations (e.g., array list, prefix-tree) and techniques for visiting the search space. Data access functions have been devised for efficiently loading in memory the index data. Once in memory, data is available for item set extraction by means of the algorithm of choice.
3. The IMine physical organization supports efficient data access during item set extraction. Correlation analysis allows us to discover data accessed together during pattern extraction. To minimize the number of physical data blocks read during the mining process, correlated information is stored in the same block.
4. IMine supports item set extraction in large data sets. We exploit a direct writing technique to avoid representing in memory the entire large data set. Direct materialization has a limited impact on the final index size because it is applied only on a reduced portion of the data set.

II. LITERATURE SURVEY

The goal of data mining is to discover important associations among items such that the presence of some items in a transaction will imply the presence of some other items. To achieve this purpose, many people propose different procedures; here we discuss some of them.

E. Baralis et al., [10] recommended itemset mining on indexed data blocks. Numerous attempts have been offered to combine data mining activities with relational DBMSs, but a correct incorporation into the relational DBMS kernel has been infrequently achieved. This paper suggested an innovative indexing method, which denotes the transactions in a succinct form, suitable for tightly incorporating frequent itemset mining in a relational DBMS. The data illustration is complete, i.e. no support threshold

Mining association rules from XML data with index table was suggested by Xin-Ye Li et al., [11]. Mining XML association rule is tackled with extra challenge because of the inherent flexibilities of XML in both arrangement and semantics. With the purpose of making mining XML association rule very efficient, this paper provides a new definition of transaction and item in XML environment, then construct transaction database depending on an index table. Based on the definition and the index

table utilized for XML searching, it is easy to check the relation among the transaction and retrieve an item quickly. A high adaptive mining approach is also illustrated. By using this approach, mining rules can be processed with no assistance of interest associations specified by users and mining unknown rules. The effectiveness of these approaches is proved with the help of experiments on real-life data.

E.J. Keogh et al., [12] proposed an indexing scheme for fast similarity search in large time series databases. This paper addresses the trouble of similarity searching in huge time-series databases. The authors proposed an innovative indexing approach that permits quicker retrieval. The index is produced by generating bins that include time series subsequences of roughly the similar shape. For every bin, this proposed approach can rapidly compute a lower bound on the distance among a given query and the most similar element of the bin. This bound permits to search the bins in greatest-first order, and to prune some bins from the search space without verifying the contents. Further speedup can be achieved by optimizing the data inside the bins in such a way that ignores the process of comparing the query to every item in the bin.

L. Golab et al., [13] proposed indexing time method for evolving data with variable lifetimes. Numerous applications store data items for a pre-determined, fixed duration of time. Examples consist of sliding windows over online data streams, in which old data are thrown out as the window slides forward. Earlier researches on management of data with limited lifetimes have emphasized online query processed in main memory. In this approach, the authors concentrate on the difficulty of indexing time-developing data on disk for offline investigation. With the intention of decreasing the I/O costs of index updates, existing work separates the data chronologically. Thus, only the previous separation is examined for expirations, only the youngest separations acquire insertions, and the remaining partitions in the middle are not processed. On the other hand, this result is based upon the hypothesis that the order in which the data are introduced is equivalent to the termination order, which means that the lifetime of each data item is the similar. In order to break this hypothesis, the authors reveal that the existing solutions no longer be relevant, and suggested a new index partitioning strategies that provide low update costs and quick access times.

A new approach of modified transaction reduction algorithm for mining frequent itemset was proposed by R.E. Thevar et al., [14]. Association rule mining is to take out the interesting association and relation among the huge volumes of transactions. This procedure is segmented into two sub problem: first problem is to discover the frequent itemsets from the

transaction and then the second problem is to build the rule from the mined frequent itemset [15]. Frequent itemsets creation is the necessary and most time huge procedure for association rule mining. Currently, most well-organized apriori-like algorithms rely deeply on the minimum support constraints to prune the enormous amount of non-candidate itemsets. These algorithms store numerous unnecessary itemsets and transactions. In this paper, the authors proposed an innovative frequent itemsets creation algorithm called MTR-FMA (modified transaction reduction depends on frequent itemset mining algorithm) that sustains its performance even at relative low supports. The experimental output also proves that proposed MTR-FMA algorithm on an outset is quicker than high efficient AprioriTid and other algorithms.

Lei Wen et al., [16] developed an efficient algorithm for mining frequent closed itemset. Association rule mining was a significant field of data mining investigation. Determining the potential frequent itemset was a vital step. The existed frequent itemset discovery algorithms could find out all the frequent itemset or maximal frequent itemset. N. Pasquier developed an innovative job of mining frequent closed itemset. The size of frequent closed itemset was much lesser than all the frequent itemsets and did not lose any information.

In this paper, we propose a new itemset approach depends on the index. This approach can discover all the frequent closed itemset powerfully by using indexing method.

III. INDEX STRUCTURE

Index structure for extracting item set as sequence of data blocks. The index supports user communication, where the user specifies many constraints for itemset extraction. It permits the mining of the complete set of itemsets which satisfy (a) time constraints and (b) support constraints. Since the index contains all feature potentially required during the mining task, the extraction can be carried out by means of the index, without accessing the database. The data representation is absolute, i.e., no support threshold is enforced

throughout the index construction stage, to permit reusing the index for mining itemsets with any support threshold. Constraints like support and confidence is not enforced throughout the index creation stage. Therefore, the extraction can be carried out using the index alone, without accessing the original database. As the databases are necessary in almost all the retail stores, super markets, etc., it is necessary to develop an approach for item set mining with the help of index support. The structure of the IMine index is characterized by two components: the Item set-Tree and the Item-Btree. The two components provide two levels of indexing. The Item set-Tree (I-Tree) is a prefix-tree which represents relation R by means of a succinct and lossless compact structure.

The Item-Btree (I-Btree) is a B+Tree structure which allows reading selected I-Tree portions during the extraction task. For each item, it stores the physical locations of all item occurrences in the I-Tree. Thus, it supports efficiently loading from the I-Tree the transactions in R including the item. In the following, we describe in more detail the I-Tree and the I-Btree structures. Fig. 1 a and b shows the complete structure of the corresponding IMine index. In the I-Tree paths (Fig. 1 a), nodes are sorted by decreasing support of the corresponding items. In the case of items with the same support, nodes are sorted by item lexicographical order. In the I-Tree, the common prefix of two transactions is represented by a single path.

A. I-Tree

The I-Tree associated to relation R is actually a forest of prefix-trees, where each tree represents a group of transactions all sharing one or more items. Each node in the I-Tree corresponds to an item in R. Each path in the I-Tree is an ordered sequence of nodes and represents one or more transactions in R. Each item in relation R is associated to one or more I-Tree nodes and each transaction in R is represented by a unique I-Tree path. Each I-Tree node is associated with a node support value, representing the number of transactions which contain (without any different interleaved item) all the items in the sub path reaching the node.

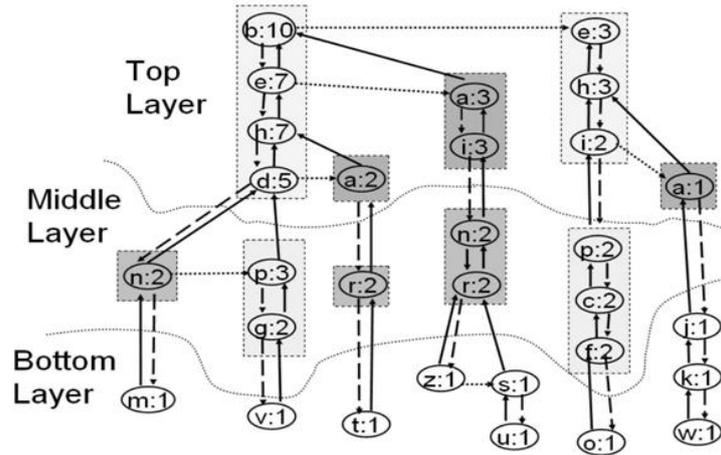


Fig 1 (a) I Tree for the example data set

B. I-Btree

The I-Btree allows selectively accessing the I-Tree disk blocks during the extraction process. It is based on a B+Tree structure [21]. Fig. 1 b shows the I-Btree for the example data set and a portion of the pointed I-Tree. For each item *i* in relation *R*, there is one entry in the I-Btree. In particular, the I-Btree

leaf associated to *i* contains *i*'s item support and pointers to all nodes in the I-Tree associated to item *i*. Each pointer stores the physical location of the record in table TI-Tree storing the node. Fig. 1 b shows the pointers to the I-Tree nodes associated to item *r*.

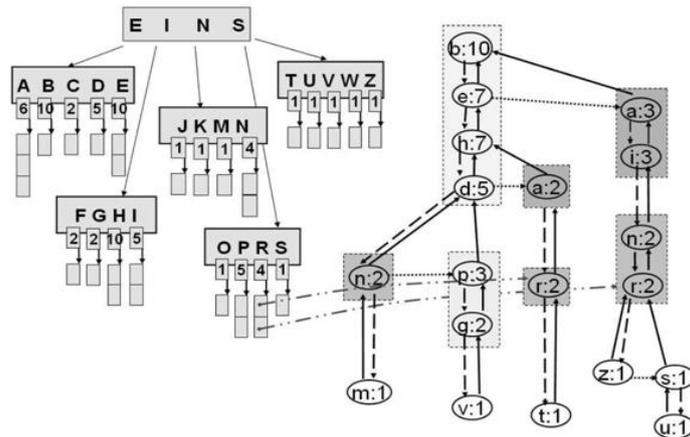


Fig 1 (b) I BTree for the example data set

IV. ITEM SET MINING

Item set mining are two sequential steps: 1) the needed index data is loaded and 2) item set extraction takes place on loaded data.

A. Frequent Item Set Extraction

This section describes how frequent item set extraction takes place on the IMine index. We present two approaches, denoted as FP-based and LCM-based algorithms, which are an adaptation of the FP-Growth algorithm [17] and LCM v.2algorithm [18], respectively.

B. FP-based algorithm:

The FP-growth algorithm stores the data in a prefix-tree structure called FP-tree. First, it computes item support. Then, for each transaction, it stores in the FP-tree its subset including frequent items. Items are considered one by one. For each

item, extraction takes place on the frequent-item projected database, which is generated from the original FP-tree and represented in a FP-tree based structure.

C. LCM-based algorithm:

The LCM v.2 algorithm loads in memory the support based projection of the original database. First, it reads the transactions to count item support. Then, for each transaction, it loads the subset including frequent items. Data are represented in memory by means of an array based data structure, on which the extraction takes place.

V. EXPERIMENTAL RESULTS

We validated our approach by means of a large set of experiments addressing the following issues:

- Performance of the IMine index creation, in terms of both creation time and index size,

- Performance of frequent item set extraction, in terms of execution time, memory usage, and I/O access time,
- Effect of the DBMS buffer cache size on hit rate,
- Effect of the index layered organization,
- Effect of direct writing, and
- Scalability of the approach.

We ran the experiments for both dense and sparse data distributions. We report experiments on six representative data sets whose characteristics (i.e., transaction and item cardinality, average

transaction size (AvgTrSz), and data set size) are in Table 1. Connect and Pumsb [19] are dense and medium-size data sets. Kosarak [19] is a large and sparse data set including click-stream data. T10I200P20D2M is a dense and large synthetic data set, while T15I100P20C1D5M and T20I100P15C1D7M are quite sparse and large synthetic data sets. Synthetic data sets are generated by means of the IBM generator [20]. For all data sets, the index has been generated without enforcing any support threshold.

Table 1. Data Set Characteristics and Corresponding Indices.

| Dataset | Dataset | | | | IMine | | |
|-----------------|--------------|--------|--------|-----------|-------------|--------------|------------|
| | Transactions | Items | AvTrSz | Size (KB) | I-Tree (KB) | I-Btree (KB) | Time (sec) |
| CONNECT | 67,557 | 129 | 43 | 25,527 | 22,634 | 4,211 | 11.05 |
| PUMSB | 98,092 | 2,144 | 37.01 | 35,829 | 57,932 | 10,789 | 34.47 |
| KOSARAK | 1,017,029 | 41,244 | 7.9 | 85,435 | 312,647 | 58,401 | 893.81 |
| T10I200P20D2M | 2,000,000 | 86,329 | 20.07 | 544,326 | 233,872 | 104,605 | 666.5 |
| T15I100P20C1D5M | 5,000,000 | 45,656 | 22 | 1,476,523 | 1,464,144 | 277,029 | 3,736.7 |
| T20I100P15C1D7M | 7,000,000 | 39,141 | 22 | 2,075,478 | 6,758,896 | 944,450 | 8350.72 |

Table 1 reports both I-Tree and I-Btree size for the six data sets. The overall IMine index size is obtained by summing both contributions. The IMine indices have been created with the default value $K_{avg}=1.2$. Furthermore, the Connect, Pumsb, Kosarak, and T10I200P20D2M data sets have been created with $K_{sup}=0$, while large synthetic data sets with $K_{sup}=0.05$. The adopted I-Tree representation is more suitable for dense data distributions, for which it provides good data compression. In dense data sets (e.g., T10I200P20D2M) where data are highly correlated, the I-Tree structure is more compact. In sparse data sets (e.g., Kosarak), where data are weakly correlated, data compression is low and storing the I-Tree requires more disk blocks.

VI. CONCLUSION AND FUTURE WORK

Due to the world wide increase in the available data, it is very difficult for obtaining the related data with better accuracy. Therefore the available techniques for data mining will not be able to extract the relevant data. This leads to the requirement for developing a better data mining technique which suits all situations. Association rule mining technique is one of the better techniques among the existing techniques. Even, this technique possesses various difficulties when large database used. Later, a new technique called indexing is introduced to solve those problems. This indexing will carry the necessary parameters to classify the required data from the large database. This paper provides object oriented approach to IMine in a complete and compact representation of transactional data. It is a general structure that efficiently supports different algorithmic approaches to item set extraction.

Experimental results show that, Selective access of the physical index blocks significantly reduces the I/O costs and efficiently exploits DBMS buffer management strategies.

REFERENCES

- [1] R. Agrawal and R. Srikant, "Fast Algorithm for Mining Association Rules," Proc. 20th Int'l Conf. Very Large Data Bases (VLDB '94), Sept. 1994.
- [2] J. Han, J. Pei, and Y. Yin, "Mining Frequent Patterns without Candidate Generation," Proc. ACM SIGMOD, 2000.
- [3] H. Toivonen, "Sampling Large Databases for Association Rules," Proc. 22nd Int'l Conf. Very Large Data Bases (VLDB '96), pp. 134-145, 1996.
- [4] M. El-Hajj and O.R. Zaiane, "Inverted Matrix: Efficient Discovery of Frequent Items in Large Datasets in the Context of Interactive Mining," Proc. Ninth ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (SIGKDD), 2003.
- [5] G. Grahne and J. Zhu, "Mining Frequent Itemsets from Secondary Memory," Proc. IEEE Int'l Conf. Data Mining (ICDM '04), pp. 91-98, 2004.
- [6] G. Ramesh, W. Maniatty, and M. Zaki, "Indexing and Data Access Methods for Database Mining," Proc. ACM SIGMOD Workshop Data Mining and Knowledge Discovery (DMKD), 2002.
- [7] Y.-L. Cheung, "Mining Frequent Itemsets without Support Threshold: With and without Item Constraints," IEEE Trans. Knowledge and Data Eng., vol. 16, no. 9, pp. 1052-1069, Sept. 2004.
- [8] G. Cong and B. Liu, "Speed-Up Iterative Frequent Itemset Mining with Constraint Changes," Proc. IEEE Int'l Conf. Data Mining (ICDM '02), pp. 107-114, 2002.
- [9] C.K.-S. Leung, L.V.S. Lakshmanan, and R.T. Ng, "Exploiting Succinct Constraints Using FP-Trees," SIGKDD Explorations Newsletter, vol. 4, no. 1, pp. 40-49, 2002.
- [10] E. Baralis, T. Cerquitelli, and S. Chiusano, "Index Support for Frequent Itemset Mining in a Relational DBMS," Proceedings 21st International Conference on Data Engineering (ICDE), pp. 754 - 765, 2005.
- [11] Xin-Ye Li, Jin-Sha Yuan and Ying-Hui Kong, "Mining Association Rules from XML Data with Index Table," International Conference on Machine Learning and Cybernetics, Vol. 7, pp. 3905 - 3910, 2007.

- [12] E.J. Keogh and M.J. Pazzani, "An indexing scheme for fast similarity search in large time series databases," Eleventh International Conference on Scientific and Statistical Database Management, pp. 56 – 67, 1999.
- [13] L. Golab, P. Prahadka and M.T. Ozsu, "Indexing Time-Evolving Data With Variable Lifetimes," 18th International Conference on Scientific and Statistical Database Management, pp. 265 – 274, 2006.
- [14] R.E. Thevar and R. Krishnamoorthy, "A new approach of modified transaction reduction algorithm for mining frequent itemset," 11th International Conference on Computer and Information Technology (ICIT 2008), pp. 1 – 6, 2008.
- [15] Jianyong Wang, J. Han, Y. Lu and P. Tzvetkov, "TFP: an efficient algorithm for mining top-k frequent closed itemsets" IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 5, pp. 652 – 663, 2005.
- [16] Lei Wen, "An efficient algorithm for mining frequent closed itemset," Fifth World Congress on Intelligent Control and Automation (WCICA 2004), Vol. 5, pp. 4296 – 4299, 2004.
- [17] J. Han, J. Pei, and Y. Yin, "Mining Frequent Patterns without Candidate Generation," Proc. ACM SIGMOD, 2000.
- [18] T. Uno, M. Kiyomi, and H. Arimura, "LCM ver. 2: Efficient Mining Algorithms for Frequent/Closed/Maximal Itemsets," Proc. IEEE ICDM Workshop Frequent Itemset Mining Implementations (FIMI), 2004.
- [19] FIMI, <http://fimi.cs.helsinki.fi/>, 2008.
- [20] N. Agrawal, T. Imielinski, and A. Swami, "Database Mining: A Performance Perspective," IEEE Trans. Knowledge and Data Eng., vol. 5, no. 6, Dec. 1993.

FAULT PREDICTION IN OBJECT-ORIENTED SYSTEMS BASED ON C³ (CONCEPTUAL COHESION OF CLASSES)

PRAKASA RAO DASARI

2/2 M.TECH CSE, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

VASANTHAKUMARI G

ASSOC.PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI
ANDHRA PRADESH, INDIA

Abstract—To remain competitive in the dynamic world of software development, organizations must optimize the usage of their limited resources to deliver quality products on time and within budget. This requires prevention of fault introduction and quick discovery and repair of residual faults. In this paper a new approach for predicting and classification of faults in object-oriented software systems is introduced. In particular, cohesion is a desirable property of software as it positively impacts understanding, reuse, and maintenance. Currently proposed measures for cohesion in Object-Oriented (OO) software reflect particular interpretations of cohesion and capture different aspects of it. Existing approaches are largely based on using the structural information from the source code, such as attribute references, in methods to measure cohesion. This paper proposes a new measure for the cohesion of classes in OO software systems based on the analysis of the unstructured information embedded in the source code, such as comments and identifiers.

Index Terms— Fault Prediction; Object Oriented Software; Conceptual Cohesion of Classes(C³).

I. INTRODUCTION

Software reliability can be defined as the probability of failure-free operation of a computer program executing in a specified environment for a specified time [1]. It is often considered a software quality factor that can aid in predicting the overall quality of a software system using standard predictive models. Predictive models of software faults use historical and current development data to make predictions about faultiness of software subsystems/modules. One of the goals of the OO analysis and design is to create a system where classes have high cohesion and there is low coupling among them. These class properties

facilitate comprehension, testing, reusability, maintainability, etc. Software cohesion can be defined as a measure of the degree to which elements of a module belong together [2]. Cohesion is also regarded from a conceptual point of view. In this view, a cohesive module is a crisp abstraction of a concept or feature from the problem domain, usually described in the requirements or specifications. Although software faults have been widely studied in both procedural and object-oriented programs, there are still many aspects of faults that remain unclear. This is true especially for object-oriented software systems, in which inheritance and polymorphism can cause a number of anomalies and fault types [3]. Unfortunately, existing techniques used to predict faults in procedural software are not generally applicable in object-oriented systems.

Proposals of measures and metrics for cohesion abound in the literature as software cohesion metrics proved to be useful in different tasks [4], including the assessment of design quality [5], [6], productivity, design, and reuse effort, prediction of software quality, fault prediction, modularization of software, and identification of reusable components [7]. Most approaches to cohesion measurement have automation as one of their goals as it is impractical to manually measure the cohesion of classes in large systems. We propose a new measure for class cohesion, named the Conceptual Cohesion of Classes (C³), which captures the conceptual aspects of class cohesion, as it measures how strongly the methods of a class relate to each other conceptually.

A. Class Cohesion

The components of a class are the instance variables and methods defined in the class plus those that are inherited. A method and an instance variable are related by the way that an instance variable is used by the method. Two methods are related (connected) through instance variable(s) if both methods use the instance variable(s). Class cohesion is defined in terms of the relative number of connected methods in the class.

B. Inheritance and Cohesion

A subclass inherits methods and instance variables from its super class. We have several options for evaluating cohesion of a subclass. We can (1) include all inherited components in the subclass in our evaluation, (2) include only methods and instance variables defined in the subclass, or (3) include inherited instance variables but not inherited methods. The class cohesion measures that we develop can be applied using any one of these options.

C. Measuring Object Oriented Reuse

We focus on private reuse within one software system [8]. We evaluate reuse from the server perspective, since this is the best orientation for evaluating reusability [9]. We are interested in two different forms of class reuse, reuse via instantiation and reuse via inheritance.

A class is reused by being instantiated in other classes or by being inherited to them. Instantiation reuse of a class is measured as the number of classes where the class is instantiated. Inheritance reuse of a class is the number of classes which inherit the class, i.e., the number of descendants (both direct and indirect descendants).

II. RELATED WORK

Software developers aim for systems with high cohesion and low coupling. The value of these goals has not been validated empirically [10]. Rather, they have been justified on the basis of intuition. The amount of reuse the number of times that a component is reused is an indicator of reusability. Of course, other factors such as the usefulness of a component are also components of reusability.

Cohesion refers to the “relatedness” of module components. A highly cohesive component is one with one basic function. It should be difficult to split a cohesive component. Cohesion can be classified using an ordinal scale that ranges from the least desirable category coincidental cohesion to the most desirable functional cohesion [11]. To apply this cohesion model to classes in object-oriented software, we need to add a new classification, data cohesion [8].

Bieman and Ott developed a set of functional cohesion measures based on program slices [12]. These measures apply only to individual functions; their application to entire classes is not obvious. Chidamber and Kemerer developed a Lack of Cohesion in Methods (LCOM) measure for object-oriented software [13]. LCOM is effective at identifying the most non-cohesive classes, but it is not effective at distinguishing between partially cohesive classes. LCOM indicates lack of cohesion only when, compared pair wise, fewer than half of the paired methods use the same instance variables.

Recently, other structural cohesion metrics have been proposed, trying to improve existing metrics by considering the effects of dependent instance variables whose values are computed from other instance variables in the class [14], [15], [16]. Other recent approaches have addressed class cohesion by considering the relationships between the attributes and methods of a class based on dependence analysis [17]. Although different from each other, all of these structural metrics capture the same aspects of cohesion, which relate to the data flow between the methods of a class.

Even though these metrics were not specifically designed for the measurement of cohesion in OO software, they could be extended to measure cohesion in OO systems.

III. AN INFORMATION RETRIEVAL APPROACH TO CLASS COHESION MEASUREMENT

OO analysis and design methods decompose the problem addressed by the software system development into classes in an attempt to control complexity. High cohesion for classes and low coupling among classes are design principles aimed at reducing the system complexity. The most desirable type of cohesion for a class is model cohesion [18] such that the class implements a single semantically meaningful concept. This is the type of cohesion that we are trying to measure in our approach.

The source code of a software system contains unstructured and (semi)structured data. The structured data is destined primarily for the parsers, while the unstructured information (that is, the comments and identifiers) is destined primarily to the human reader. Our approach is based on the premise that the unstructured information embedded in the source code reflects, to a reasonable degree, the concepts of the problem and solution domains of the software, as well as the computational logic of the source code. This information captures the domain semantics of the software and adds a new layer of semantic information to the source code, in addition to the programming language semantics.

A. An example of measuring C3

To better understand the C3 metric, consider a class $c \in C$ with five methods $m1, m2, m3, m4, m5$. The conceptual similarities between the methods in the class are shown in Table 1. For the computation of ACSM we consider all pairs of different methods, thus $ACSM(c) = 0.5$. Since the value is positive, $C3(c) = ACSM(c) = 0.5$. This particular value for C3 does not indicate high cohesion for class c nor a low one, but the CSM values from Table 1 show that $m1$ and $m3, m2$ and $m4, m2$ and $m5$, and $m4$ and

$m5$ are closely related respectively (i.e., the CSM between each pair is larger than C3). As one can see in this example, CSM is not a transitive measure. Since C3 is an average measure, we could have situations when some pairs of methods are highly related and other are not and the average is around 0.5. With that in mind, we refine the C3 to measure the influence of the difference between the highly related and unrelated pairs of methods on the cohesion of the class.

Table 1. Conceptual similarities between the methods in class c . $ACSM(c) = 0.5$.

| | m1 | m2 | m3 | m4 | m5 |
|----|----|------|------|------|------|
| m1 | 1 | 0.21 | 0.72 | 0.33 | 0.42 |
| m2 | | 1 | 0.28 | 0.91 | 0.66 |
| m3 | | | 1 | 0.37 | 0.27 |
| m4 | | | | 1 | 0.89 |
| m5 | | | | | 1 |

B. An example of measuring LCSM

Consider the same class c described in section 3.3 with $C3(c) = 0.5$. For each method of the class c , we compute M_i based on definition 4: $M1 = \{m3\}$, $M2 = \{m4, m5\}$, $M3 = \{m1\}$, $M4 = \{m2, m5\}$, $M5 = \{m2, m4\}$. Table 1 shows us the intersection among

all pairs of sets $M_i \cap M_j$ in class c . Based on the intersection $P = \{(M1, M2); (M1, M3); (M1, M4); (M1, M5); (M2, M3); (M3, M4); (M3, M5)\}$ and $|P| = 7$. $Q = \{(M2, M4); (M2, M5); (M4, M5)\}$ and $|Q| = 3$. Thus, $LCSM(c) = 7-3 = 4$.

Table 2. Intersection results for method sets

| | M1 | M2 | M3 | M4 | M5 |
|----|----|-------------|-------------|-------------|-------------|
| M1 | | \emptyset | \emptyset | \emptyset | \emptyset |
| M2 | | | \emptyset | m5 | m4 |
| M3 | | | | \emptyset | \emptyset |
| M4 | | | | | m4 |
| M5 | | | | | |

The two results combined indicate a lower value for the cohesion of class c from the example. In another situation, class c' could have had more highly related methods than in this case (i.e., four pairs) and less unrelated method pairs with the same $C3(c')$ value (i.e., 0.5). Assume Table 2 would indicate 6 pairs of method sets with non empty intersection and only 4 with an empty intersection. The $LCSM(c')$ in that case would be 0. The combined measures will indicate that c' is more cohesive than c .

IV. SYSTEM ARCHITECTURE

The measuring methodology for the proposed cohesion metrics is described in Figure 1. The following steps are necessary to compute the C3 and LCSM metrics:

- Preprocessing and parsing of the source code to produce a text corpus. Comments and identifiers from each method are extracted and processed. A document in the corpus is created for each method in every class.
- An IR method is used to index the corpus and create an equivalent semantic space.

- Based on the IR indexing conceptual similarities are computed between each pair of methods.

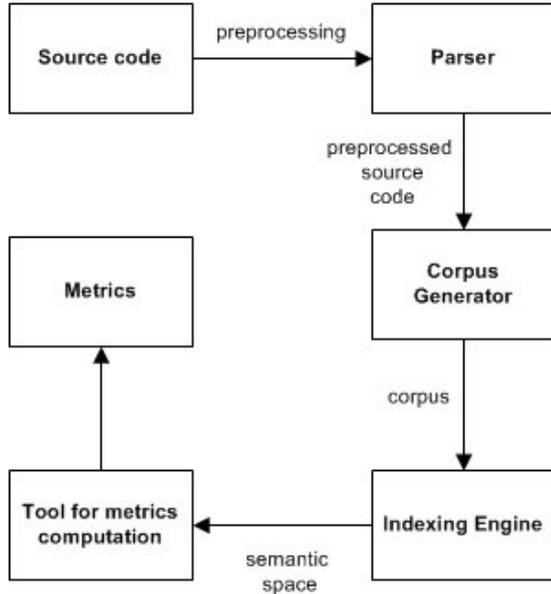


Fig 1. Measuring methodology and tools

We implemented a tool to compute C3 and LCSM for C++ software projects in MS Visual Studio .NET, based on the above methodology. Our source code parser component is based on the “Visual C++ Object Extensibility Model”. Using project information retrieved from Visual Studio .NET, the tool retrieves parts of source code that are used to produce a corpus. The extracted comments and identifier are processed by elimination of stop words and splitting identifiers that follow predefined coding standards. The corpus is indexed by the indexing engine, which is an implementation of LSI. We use the cosine between vectors in the LSI space to compute conceptual relations.

V. IMPLEMENTATION

We developed a tool IR-based Conceptual Cohesion Class Measurement, which supports this methodology and automatically computes C3 for any class in a given software system. The following steps are necessary to compute the C3 metric:

A. Corpus creation

The source code is preprocessed and parsed to produce a text corpus. Comments and identifiers from each method are extracted and processed. A document in the corpus is created for each method in every class.

B. Corpus indexing

LSI is used to index the corpus and create an equivalent semantic space. Computing conceptual similarities. Conceptual similarities are computed between each pair of methods.

- Based on the conceptual similarity measures, C3 and LCSM are computed for each class.

C. Computing C3

Based on the conceptual similarity measures, C3 is computed for each class (definitions are presented in the next section). IRC3Mis implemented as an MS Visual Studio .NET addin and computes the C3 metric for C++ software projects in Visual Studio based on the above methodology. Our source code parser component is based on the Visual C++ Object Extensibility Model. Using project information retrieved from Visual Studio .NET, the tool retrieves parts of the source code that are used to produce a corpus. For software projects that are developed outside the .NET environment, that is, Mozilla from our case study, we use external parsers and a set of our own utilities to construct the corpus. The extracted comments and identifiers are processed in the elimination of stop words and splitting identifiers that follow predefined coding standards. We use the cosine between vectors in the LSI space to compute conceptual relations. A Java version of the tool is being developed as an Eclipse plug-in.

VI. RESULTS

First, we performed the univariate logistic regression. The R² coefficient is defined as the proportion of the total variation in the dependant variable y (the fault proneness of a class) that is explained by the regression model. The bigger the value of R², the larger the portion of the total variance in y that is explained by the regression model and the better the dependent variable y is explained by the explanatory variables.

In order to evaluate logistic regression models based on the studied metrics and their combinations, we utilize the following quantitative characteristics: precision, correctness, and completeness. We use these measures to be consistent with previously published results [19, 20]. Note that these characteristics of the results are somewhat different from the precision and recall measures used in IR.

Precision here is used to evaluate how well the model classifies faulty and nonfaulty classes. For example, C3 used as a separate explanatory variable in the univariate logistic model classified 1,267 (667 as nonfaulty + 600 as faulty) classes correctly out of 2,042 classes for Mozilla, that is, a precision of 62.05 percent (see Table 4). The results of the univariate logistic regression indicate that the model based on C3 is better than any other model except that of LCOM3.

Correctness is used to show what percentage of the faulty predicted classes is really faulty (computed as the number of classes observed

and predicted as faulty divided by the total number of classes predicted as faulty). In the case of the univariate logistic regression model based on C3, the correctness is 61.35 percent since it suggested

978 classes as containing faults, but, in fact, only 600 of those have faults. Fig 2,34 shows the C out put file, C++ out put file and JAVA out put file respectively.

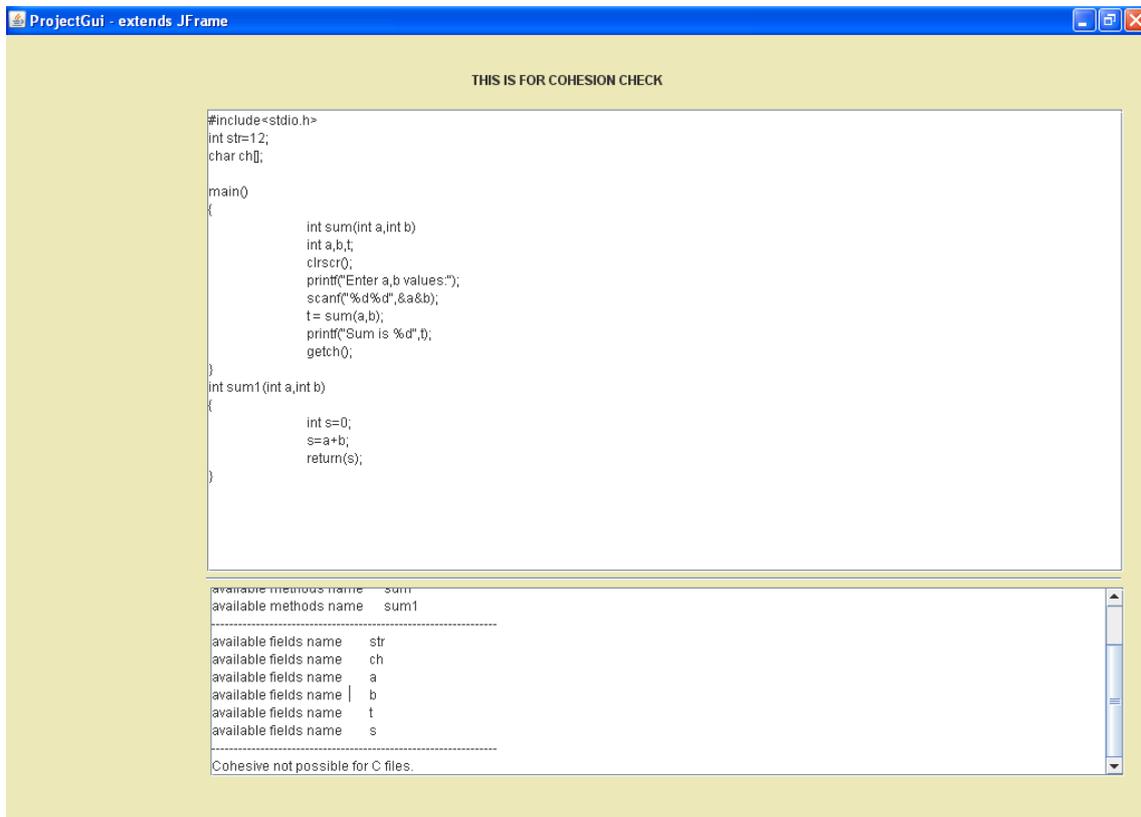


Fig 3 C OUTPUT File

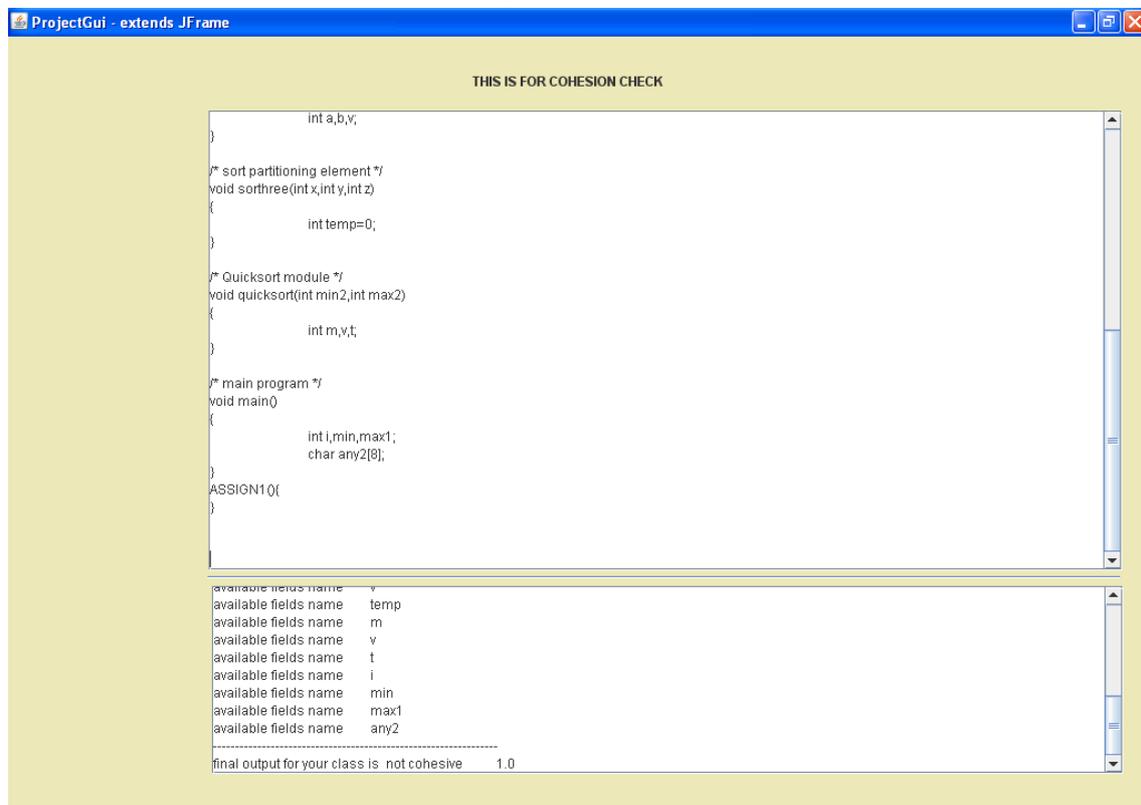


Fig 4 C++ OUTPUT File

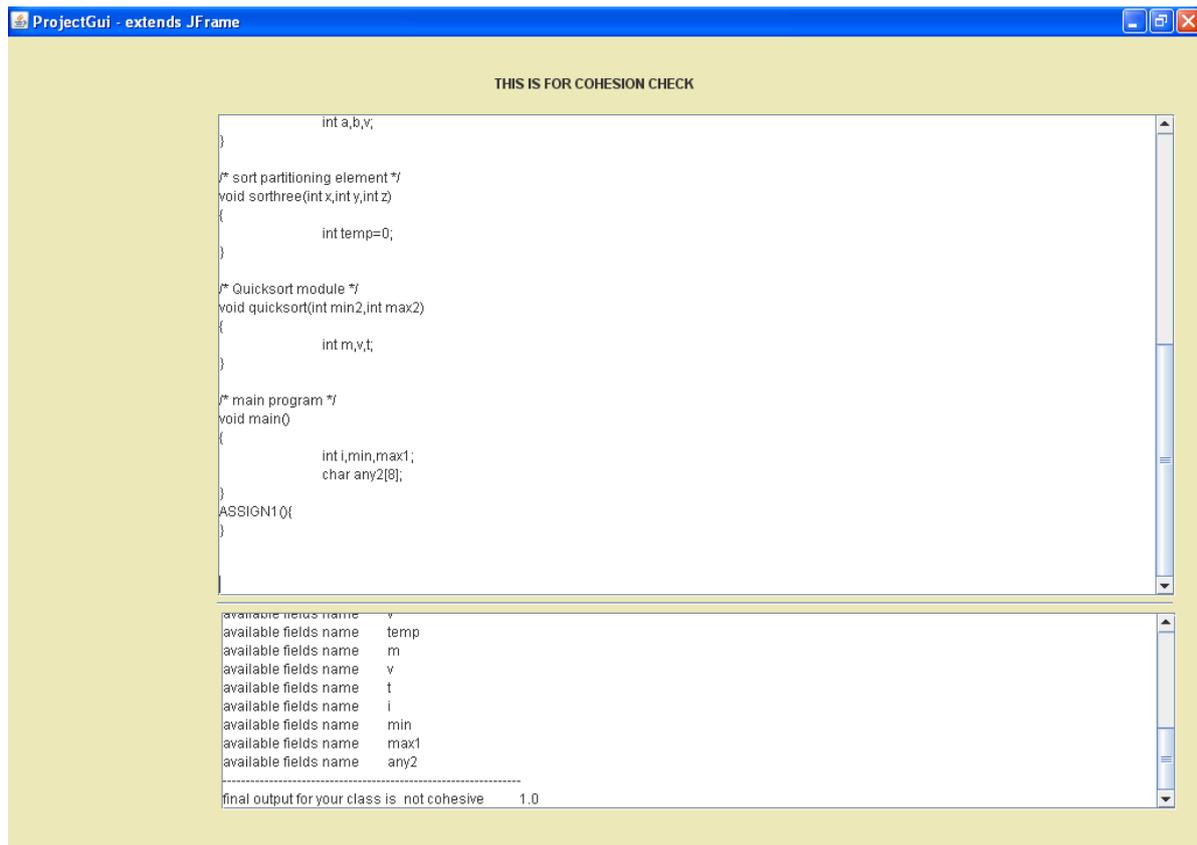


Fig 5 JAVA OUTPUT File

VII. CONCLUSIONS

Classes in object-oriented systems, written in different programming languages, contain identifiers and comments which reflect concepts from the domain of the software system. This information can be used to measure the cohesion of software. To extract this information for cohesion measurement, Latent Semantic Indexing can be used in a manner similar to measuring the coherence of natural language texts. Our results show that the classes that are heavily reused via inheritance exhibit lower cohesion. We expected to find that the most reused classes would be the most cohesive ones. Studies of additional software systems are needed to confirm these results.

Future Work

The C3 metric depends on reasonable naming conventions for identifiers and relevant comments contained in the source code. When these are missing, the only hope for measuring any aspects of cohesion rests on the structural metrics. In addition, methods such as constructors, destructors, and accessory may artificially increase or decrease the cohesion of a class. Although we did not exclude them in the results presented here, our method may be extended to exclude them from the computation of the cohesion by using approaches for identifying types of method stereotypes. C3 does not take into

account polymorphism and inheritance in its current form. It only considers methods of a class that are implemented or overloaded in the class. One way in which we can extend our work to address inheritance when building a corpus is to follow the approach in, where the source code of inherited methods is included into the documents of derived classes.

References

- [1] J. D. Musa, A. Iannino, and K. Okumoto, *Software Reliability Measurement, Prediction, Application*. the United States of America: McGraw- Hill Book Company, 1987.
- [2] J. Bieman and B.-K. Kang, "Cohesion and Reuse in an Object-Oriented System," *Proc. Symp. Software Reusability*, pp. 259-262, Apr. 1995.
- [3] J. Offutt and R. Alexander, "A fault model for subtype inheritance and polymorphism," in *12th International Symposium on Software Reliability Engineering*, November 2001, pp. 84 – 95.
- [4] D. Darcy and C. Kemerer, "OO Metrics in Practice," *IEEE Software*, vol. 22, no. 6, pp. 17-19, Nov./Dec. 2005.
- [5] J. Bansiya and C.G. Davis, "A Hierarchical Model for Object-Oriented Design Quality Assessment," *IEEE Trans. Software Eng.*, vol. 28, no. 1, pp. 4-17, Jan. 2002.
- [6] L.C. Briand, J. Wu, J.W. Daly, and V.D. Porter, "Exploring the Relationship between Design Measures and Software Quality in Object-Oriented Systems," *J. System and Software*, vol. 51, no. 3, pp. 245-273, May 2000.
- [7] J.K. Lee, S.J. Jung, S.D. Kim, W.H. Jang, and D.H. Ham, "Component Identification Method with Coupling and Cohesion," *Proc. Eighth Asia-Pacific Software Eng. Conf.*, pp. 79-86, Dec. 2001.
- [8] N. Fenton, *Software Metrics - A Rigorous Approach*. Chapman and Hall, London, 1991.
- [9] J. Bieman, "Deriving measures of software reuse in object-oriented systems." *Proc. BCS-FA CS Workshop on*

- Formal Aspects of Measurementj pp. 79–82. Springer-Verlag, 1992.
- [10] N. Fenton, S.L. Pfleeger, and R. Glass. Science and substance: a challenge to software engineers. *IEEE Sofiware*, 11(4):86-95, July 1994.
- [11] E. Yourdon and L. Constantine. Prentice-Hall, Englewood Cliffs, Structured Design.NJ, 1979.
- [12] J. Bieman and L. Ott. Measuring functional cohesion. *IEEE Trans. Software Engineering*, 20(8) :644–657, Aug. 1994.
- [13] S. Chidamber and C. Kemerer. A metrics suite for object oriented design. *IEEE Trans. Software Engineering*, 20(6):476–493, June 1994.
- [14] H.S. Chae, Y.R. Kwon, and D.H. Bae, “Improving Cohesion Metrics for Classes by Considering Dependent Instance Variables,” *IEEE Trans. Software Eng.*, vol. 30, no. 11, pp. 826-832, Nov. 2004.
- [15] Y. Zhou, L. Wen, J. Wang, Y. Chen, H. Lu, and B. Xu, “DRC: A Dependence-Relationships-Based Cohesion Measure for Classes,” *Proc. 10th Asia-Pacific Software Eng. Conf.*, pp. 215-223, 2003.
- [16] Y. Zhou, B. Xu, J. Zhao, and H. Yang, “ICBMC: An Improved Cohesion Measure for Classes,” *Proc. 18th IEEE Int’l Conf. Software Maintenance*, pp. 44-53, Oct. 2002.
- [17] Z. Chen, Y. Zhou, B. Xu, J. Zhao, and H. Yang, “A Novel Approach to Measuring Class Cohesion Based on Dependence Analysis,” *Proc. 18th IEEE Int’l Conf. Software Maintenance*, pp. 377- 384, 2002.
- [18] J. Eder, G. Kappel, and M. Schreft, “Coupling and Cohesion in Object-Oriented Systems,” technical report, Univ. of Klagenfurt, 1994.
- [19] V.R. Basili, L.C. Briand, and W.L. Melo, “A Validation of Object- Oriented Design Metrics as Quality Indicators,” *IEEE Trans. Software Eng.*, vol. 22, no. 10, pp. 751-761, Oct. 1996.
- [20] T. Gyimo’thy, R. Ferenc, and I. Siket, “Empirical Validation of Object-Oriented Metrics on Open Source Software for Fault Prediction,” *IEEE Trans. Software Eng.*, vol. 31, no. 10, pp. 897-910, Oct. 2005.
-

STATIC ANALYSIS TOOL FOR DETECTING WEB APPLICATION VULNERABILITIES

L. VENKATA SATYANARAYANA,

2/2 M.TECH CSE, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI,
ANDHRA PRADESH, INDIA.

M.V.B.CHANDRA SEKHAR,

ASSOCIATE PROFESSOR,
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
ADITYA INSTITUTE OF TECHNOLOGY AND MANAGEMENT, TEKKALI,
ANDHRA PRADESH, INDIA.

Abstract— Over the past few years, injection vulnerabilities have become the primary target for remote exploits. SQL injection, command injection, and cross-site scripting are some of the popular attacks that exploit these vulnerabilities. Many web applications written in ASP suffer from injection vulnerabilities, and static analysis makes it possible to track down these vulnerabilities before they are exposed on the web. In this paper we propose a new technique to detect XSS attacks and SQL injection vulnerabilities based on taint analysis, It tracks various kinds of external input, tags taint types, constructing control flow graph is constructed based on the use of data flow analysis of the relevant information, taint data propagate to various kinds of vulnerability functions, and detect the XSS or SQL Injection vulnerability in web application's source code. Results show the benefits of the tool in identifying potential security vulnerabilities.

Index Terms— W3; SQL Injection; XSS.

I. INTRODUCTION

Web applications are growing more and more popular as both the availability and speed of the internet increase. Many web servers nowadays are equipped with some sort of scripting environment for the deployment of dynamic web applications. However, most public web hosting services do not enforce any kind of quality assurance on the applications that they run, which can leave a web server open to attacks from the outside[1].

Poorly written web applications are highly vulnerable to attacks because of their easy accessibility on the internet. One careless line of program code could potentially bring down an entire computer network. The reasons for the increase of threats in Web application [2,3] could be

divided into two main parts: On one hand, software are developing in too large a scale together with the expanding complexity and extensibility of software while flaws still exist in their source codes; On the other hand, This is probably due to ease of detection and exploitation of web vulnerabilities, combined with the proliferation of low-grade software applications. At the moment, the overflow of Web application programs and Plug-in lead to the result that much of the code is alpha or beta, written by inexperienced programmers with easy-to learn languages such as ASP (Active Server Pages).

Such software is often rife with easy-to-find vulnerabilities, even malicious hideaway back door. For instance, injection threats exist in the early version of Eweb editor and the fckeditor. Security problem in software refers to threats incurred because of the flaws in software research, designation, programming, testing and implementation [4]. They are taken use of by attackers so as to change the function of the software from original intention of the software designers. As a typical Web application attacks, the most popular is the SQL injection and XSS, because the most basic data manipulations for these vulnerabilities are very simple to perform, e.g. ''' for SQL injection and '<script> alert('hi') </script>' for XSS [5]. This makes it easy for beginning researchers to quickly test large amounts of software.

II. RELATED WORK

Static analysis security tools attempt to find security vulnerabilities without executing the software by scanning the source code for known potentially security-compromising functions[6]. They then perform analyses to try to determine if, indeed, a function call could be maliciously attacked. These

tools cannot guarantee to find all security vulnerabilities in a program and often report many false positives (those potential vulnerabilities reported by a tool which are not actual vulnerabilities).

In 2005 and 2006, XSS was number 1, and SQL injection was number 2 [7]. In 2009 around domestic college entrance examination online enrollment, the domestic university's admission websites suffer the threat by "Trojan horse" attack, when the user visits the page that attacked by XSS, the user's sensitive data are stolen. These explained that strengthens the Web application security the work to be urgent. ITS4 was one of the first available static security analysis tools to search C source code looking for potentially dangerous function calls [8]. ITS4 performs limited analysis to determine how risky a function call is and, for every problem reported provides suggestions how to mitigate the security vulnerability. RATS is similar to ITS4 in its approach but performs additional analysis to attempt to reduce the number of false positives reported [9]. Unlike ITS4, however, RATS performs analysis to discover Time Of Check, Time Of Use race conditions.

Splint (Secure Programming Lint) is an improvement over another static security analysis tool, Lint [10] that does additional analysis on potential security vulnerabilities beyond both ITS4 and RATS. Other tools perform different analysis techniques to try and discover a different type of security vulnerability or eliminate a different type of false positives. For example, BOON [11] performs analysis focusing primarily on the detection of the buffer overflow security vulnerability whereas Flaw Finder [12] uses a vulnerability database as does ITS4 and RATS. Thus, different tools often produce different sets of results.

III. SYSTEM ARCHITECTURE

A web application, as the name implies, is a computer application that is accessed through a web-based user interface. This is typically implemented through a client-server setup, with the server running an HTTP server software package (such as Apache or Microsoft IIS) capable of generating dynamic web pages, while the client communicates with the server through a web browser (such as Microsoft Internet Explorer or Mozilla Firefox). The working of such a web application is roughly sketched in Figure 1.

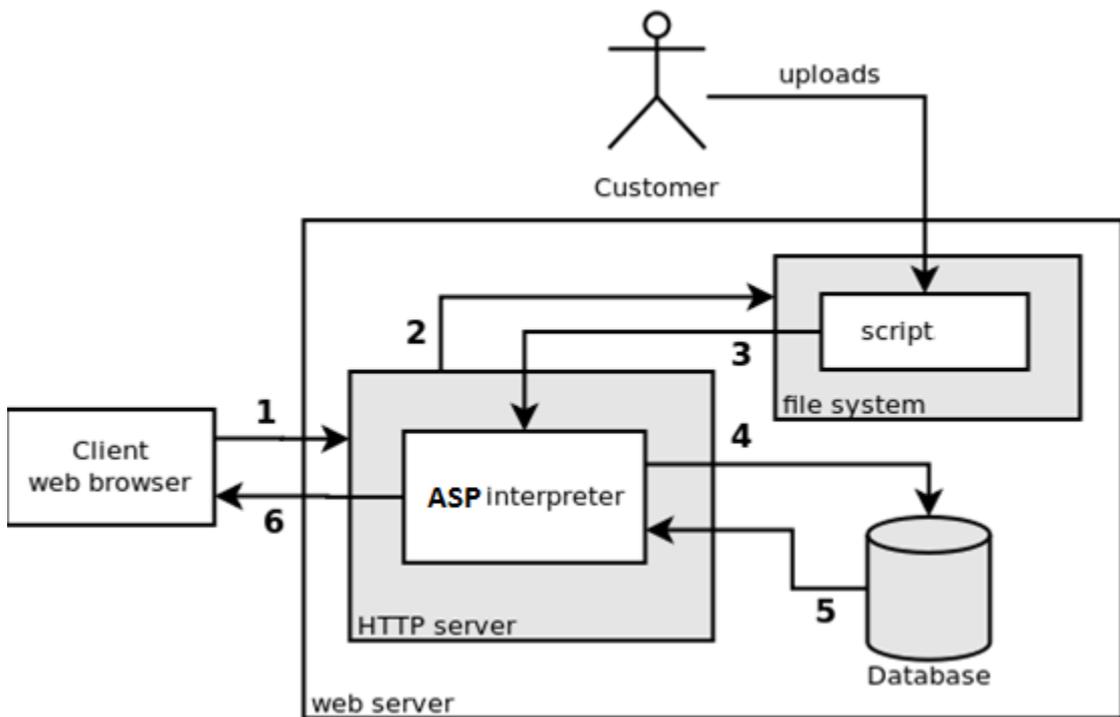


Fig 1 System architecture

Whenever the client interacts with the server, communication takes place in the form of HTTP requests. The client sends a request (typically a GET or POST request) to the server, along with a series of parameters (step 1 in Fig. 1). The HTTP server recognizes that this is a request for a dynamic

web page, in this case a page that is to be generated by a ASP script. It fetches the corresponding ASP script from the web server's file system (step 2) and sends it off to be processed by the integrated ASP interpreter (step 3). The ASP interpreter then executes the ASP script, making use of external

resources whenever necessary. External resources can for example be a database (steps 4 and 5), but also the file system or an API exposed by the operating system. The executed ASP script typically produces output in the form of an HTML page, which is sent back to the client and displayed in the web browser (step 6).

The most prevalent and most exploited vulnerabilities in web applications are cross-site scripting (XSS) and SQL injection (SQLI). According to a top ten composed by the Open Web Application Security Project (OWASP), XSS and SQLI were the top two most serious web application security flaws for both 2007 and 2010 [13]. According to this list, the top five of security flaws has not changed over the past three years. Vulnerabilities occurring in real-world applications are much more complicated and subtle than the examples appearing in this section. In many cases, user data is utilized in applications in ways that appear to be safe on the surface. However, due to complex interactions that is difficult to predict, unsafe data can still slip through in specific edge cases. Such vulnerabilities are hard to spot, even when using professional coding standards, careful code reviewing, and extensive testing.

A. Cross-Site Scripting

Cross-site scripting (XSS) is a type of vulnerability that allows attackers to inject unauthorized code into a web page, which is interpreted and executed by the user's web browser. XSS has been the number one web application vulnerability for many years, and according to White Hat Security, has been responsible for 66% of all website attacks in 2009 [14].

Web pages can include dynamic code written in JavaScript to allow the web page's content to be altered within the web browser as the user interacts with it. Normally, a web browser will only execute JavaScript code that originates from the same domain as the web page itself, and that code is only executed within a self-contained sandbox environment. This is the so-called Same Origin Policy [15]. This policy prevents attackers from making web browsers execute un-trusted code from an arbitrary location.

| | |
|---|--|
| 1 | <html> |
| 2 | <body> |
| 3 | <% |
| 4 | 'Re t r i e v e the user ' s name from a |
| 5 | form |
| 6 | name = Request.Response('name') |
| 7 | // Pr int the user ' s name back to them |
| 8 | Response.write "Hello there , \$name! |
| 9 | How are you doing?" |
| | %> |
| | </body> |

| |
|---------|
| </html> |
|---------|

Fig 2. Example of an XSS vulnerability

B. SQL Injection

SQL injection is a taint-style vulnerability, whereby an unsafe call to a database is abused to perform operations on the database that were not intended by the programmer. White Hat Security's report for 2009 lists SQL injection as responsible for 18% of all web attacks, but mentions that they are under-represented in this list because SQL injection flaws can be difficult to detect in scans [14].

SQL, or Structured Query Language, is a computer language specially designed to store and retrieve data in a database. Most database systems (e.g. MySQL, Oracle, Microsoft SQL Server, SQLite) use a dialect of SQL as a method to interact with the contents of the database. Scripting languages such as PHP o_er an interface for programmers to dynamically construct and execute SQL queries on a database from within their program.

It is common practice for programmers to construct dynamic database queries by means of string concatenation. Listing 2 shows an example of an SQL query that is constructed and executed from a PHP script using this method. The variable \$username is copied directly from the input supplied by the user and is pasted into the SQL query without modifications. Although this query is constructed and executed on the server and users can not directly see the vulnerable code, it is possible through a few assumptions and educated guesses to find out that there is a database table called users and that the entered user name

The test method of common SQL Injection attack is the use "'", "union", "--;" and so on key words, in test dynamic SQL sentence in program whether to exist injection vulnerability. For example, consider the login page of a web application that expects a user-name and the corresponding password. When the credentials are submitted, they are inserted within a query template such as the following:

```
“select * from admin where username =” + request.form(“username”) + “ and Password = ” + request.form(“passwd”) + “”
```

Instead of a valid user name, the malicious user sets the “username” variable to the string: ‘ or 1=1; - -’, causing the Vbscript to submit the following SQL query to the database:

```
“select * from admin where username = ‘ or 1=1 ; - - ‘ and Password = ‘ any_passwd’ ”
```

Therefore, the password value is irrelevant and may be set to any character string. The result set of the query contains at least one record, since the “where” clause evaluates to true. If the application

identifies a valid user by testing whether the result set is non-empty, the attacker can bypass the security check.

IV. STATIC ANALYSIS METHODS

Detection software security vulnerabilities are mainly dynamic analysis, formal method validation and static analysis. Static analysis is divided as type inference, data flow analysis and constraints analysis [1,16,17].

A. Type Inference

Type system of programming language concludes type definition and rules for type equivalence, type inclusiveness and type dedication. Type dedication is to derive the types of variables and methods within a program automatically so as to determine whether or not their visit meet these type rules. This kind of dedication could be used to examine the bug in types and conduct necessary type transmission with proper operations. It boasts the characteristics of simplicity and high efficiency which makes it perfect for quick detection of security threats in software. Now it is mainly applied in detection of format string vulnerability, OS kernel vulnerable pointer use.

B. Data-Flow Analysis

Data-flow analysis is used in the process programming, which collect semantic information from programs and then define and use the variables with algebraic approach. It is used in program optimization, program validation, debugging, parallel, Vectorization and serial program environment. Its realization makes use of the pair "variable definition-quoting".

C. Constraint Analysis

Constraint analysis divides program analysis into constraint generation and constraint solution. The former constructs variable type with constraint generation rules or analyses constraint system among statuses. While the later solve such constraint systems. Constraint system is comprised of equation constraint, set constraint and incorporate constraint. In the first kind, only equation exists between constraint objects. Set constraint takes program variables as a set of values, whose evaluation is regarded as conclusion relation between set expressions? While the last constraint concludes equation constraint part and set constraint part.

D. The Comparison among three Main methods

The three main methods mentioned are all explain the abstract semantics of programs and construct mathematic models based on the program property,

with which they determine the property of the program. In comparison, constraint analysis boasts the greatest ability in detection while the lowest speed of that, which makes it fit for security examination of software, data-flow analysis has relatively high speed and remarkable ability of detection which is appropriate in static analysis which should take control flow information and requires only simple operation among variable properties; when it comes to type dedication, it has the poorest ability and the fastest speed in examination and suits for security test in finite property domain and unrelated control flow.

The website that issued news or BBS forum is one kind of web application. Analyzing the logic characteristics of its services, it is not complex to find out that the process of dataflow could be give a summaries: data input (parameters) → data service processing (web server)→result output (HTML).

Based on principles of XSS threats and SQL injection threats, we could see vulnerability is mainly generated from the sanitation process of input data. Thus sanitation process of all the input data would neglect such vulnerability of taint data (outside client input data). In this paper, we would examine the code with data flow analysis; the function framework of system is shown in fig 2.

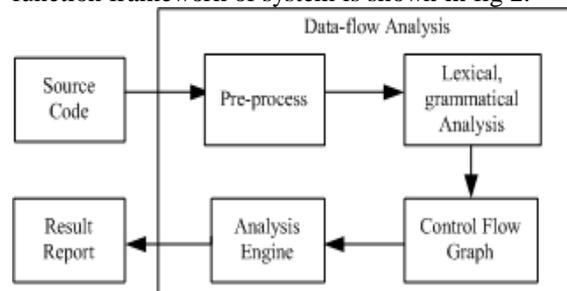


Fig 3. Function Framework of Code Review System

V. RESULTS

To test the validity of our approach, we select three open source program written by ASP. These software are commonly used a source codes of tools in network application layer and are representative. Test environment: Intel XEON CPU: 3.0GHz, 1GB cache, Windows 2003 Server , IIS6.0. Then we conduct penetrating examination in Acunetix Web with the results shown in Fig 4. We develop a tools named as ASPWC. The number of XSS reported by Acunetix Web tools is success of the XSS attack test. Possible SQL Injection vulnerability through data that is read from the Request object without any input validation. These warnings are very likely bugs that must be fixed. Sample Web Page is shown in Fig 4

```

samplepage.asp - Notepad
<!DOCTYPE html>
<HTML>
<HEAD>
<META NAME="GENERATOR" Content="Microsoft Visual Studio 6.0">
</HEAD>
<BODY>
<%
Err.Clear
ON ERROR RESUME NEXT
strAuthor = Request.Form("AUTHORNAME")
If strAuthor = "" Then
    Response.Write "AUTHORNAME is not supplied"
    Response.End
End If

strConnectString = "Provider=SQLEDB.1; Data Source=sqlmac;Initial Catalog=Test;Integrated Security=SSPI;"
Set objConn = Server.CreateObject("ADODB.CONNECTION")
Set objRS = Server.CreateObject("ADODB.RECORDSET")
Set objCommand = Server.CreateObject("ADODB.COMMAND")

' Connect to SQL Server
objConn.Open(strConnectString)

If Err.number Then
    ' Log the error and transfer control to a different page.
    Server.Transfer ("Maintenance.asp")
End If

' Execute the command
strCmd = "select title, description from books where author_name = " & strAuthor & ""
Set objCommand.ActiveConnection = objConn
objCommand.CommandText = strCmd
objCommand.CommandType = adCmdText

Set objRS = objCommand.Execute()

' Process the resultset
Do Until objRS.EOF
    Response.Write Server.HtmlEncode(objRS.Fields("title")) & vbCRLF
objRS.MoveNext
    
```

Fig 4 Sample web page

```

C:\WINDOWS\system32\cmd.exe
C:\Documents and Settings\user\Desktop\mtech\satya>msscasi_asp.exe /input="samplepage.asp" /NoLogo
samplepage.asp(37) : warning C80400: Unvalidated HTTP request data possibly executed, making 'VBSMAI potentially vulnerable to first-order SQL Injection attack'
object OBJCOMMAND (created as return.FORM 11).

Path summary:
- {return.FORM}[return.FORM`11 : string_unvalidated] created on 'Request' (line 11)
- {return.FORM}[return.FORM`11 : string_unvalidated] to {STRAUTHOR, return.FORM}[return.FORM`11 : string_unvalidated] by assignment (line 11)
- {STRAUTHOR, return.FORM}[return.FORM`11 : string_unvalidated] to {STRAUTHOR, STRCMD, return.FORM}[return.FORM`11 : string_unvalidated] on 'Transfer' (line 32)
- {STRAUTHOR, STRCMD, return.FORM}[return.FORM`11 : string_unvalidated] to {OBJCOMMAND}[return.FORM`11 : command_unvalidated] on 'TaintCommand' (line 34)
- {OBJCOMMAND}[return.FORM`11 : command_unvalidated] to {OBJCOMMAND}[return.FORM`11 : $error] on 'Execute' (line 37)
: Lines: 9, 10, 11, 13, 18, 19, 20, 21, 24, 26, 32, 33, 34, 35, 37

C:\Documents and Settings\user\Desktop\mtech\satya>
    
```

Fig 5 Output from our tool

Comments: strAuthor is assigned a value from Request.QueryString("AUTHORNAME") on line number 11 and is eventually used in the construction of dynamic SQL and executed through OBJCOMMAND on line number 37
 Use parameterized SQL query to mitigate the SQL Injection identified by the tool

| Table 1 Sample Code |
|--|
| <pre> ' Execute the command strCmd = "select title, description from books where author_name = ?" Set objCommand.ActiveConnection = objConn objCommand.CommandText = strCmd objCommand.CommandType = adCmdText Set param1 = objCommand.CreateParameter ("author", adWChar, adParamInput, 50) param1.value = strAuthor objCommand.Parameters.Append param1 </pre> |

Possible SQL Injection vulnerability through data that is read from the Request object where the input is passed through some unknown function calls that might perform data validation. If there is no data validation done inside the function call, then these are likely bugs else these are likely false positives

```

samplepage2.asp - Notepad
File Edit Format View Help
<? Language=VBScript %>
<HTML><BODY>
<%
Err.Clear : ON ERROR RESUME NEXT
strAuthor = Request.Form("AUTHORNAME")
IF Not ValidateInput(strAuthor) Then
    Server.Transfer ("Errorpage.asp")
End IF
strConnectionString = "Provider=SQLEDB.1; Data Source=sqlmac;Initial Catalog=Test;Integrated Security=SSPI;";
Set objConn = Server.CreateObject("ADODB.CONNECTION")
Set objRS = Server.CreateObject("ADODB.RECORDSET")
Set objCommand = Server.CreateObject("ADODB.COMMAND")
objConn.Open(strConnectionString)
IF Err.number Then
    'Log the error and transfer control to a different page.
    Server.Transfer ("Maintenance.asp")
End IF
Execute the command
strCmd = "select title, description from books where author_name = '' & strAuthor & '"
Set objCommand.ActiveConnection = objConn
objCommand.CommandText = strCmd
objCommand.CommandType = adCmdText
Set objRS = objCommand.Execute()
Do Until objRS.EOF
    Response.Write Server.HtmlEncode(objRS.Fields("title")) & vbCRLF
    objRS.MoveNext
Loop
objRS.Close
Set objRS = Nothing
objConn.Close
Set objConn = Nothing
Private Function ValidateInput(ByVal strInput)
    ValidateInput = true
    Set reg = New RegExp
    reg.Global = True
    reg.Pattern = "[A-Za-z0-9]+$"
    IF Not reg.Test(strAuthor) Then
        'Accept only valid input and reject all other input
        ValidateInput = False
    End IF
End Function
%>

```

Fig 6 Sample web page

```

C:\WINDOWS\system32\cmd.exe
Path summary:
- {return.FORM}[return.FORM`5 : string_unvalidated] created on 'Request' (line 5)
- {return.FORM}[return.FORM`5 : string_unvalidated] to {STRAUTHOR, return.FORM}[return.FORM`5 : string_unvalidated] by assignment (line 5)
- {STRAUTHOR, return.FORM}[return.FORM`5 : string_unvalidated] to {STRAUTHOR, return.FORM}[return.FORM`5 : string_unvalidated_low] on 'unknownCall' (line 6)
- {STRAUTHOR, return.FORM}[return.FORM`5 : string_unvalidated_low] to {STRAUTHOR, STRCMD, return.FORM}[return.FORM`5 : string_unvalidated_low] on 'Transfer' (line 19)
- {STRAUTHOR, STRCMD, return.FORM}[return.FORM`5 : string_unvalidated_low] to {OBJCOMMAND}[return.FORM`5 : command_unvalidated_low] on 'TaintCommand' (line 21)
- {OBJCOMMAND}[return.FORM`5 : command_unvalidated_low] to {OBJCOMMAND}[return.FORM`5 : $error] on 'Execute' (line 23)
: Lines: 4, 5, 6, 9, 10, 11, 12, 13, 14, 19, 20, 21, 23

C:\Documents and Settings\user\Desktop\mtech\satya>

```

Fig 7 Output from our tool

The experimental results analysis: experimental data can be seen from the above, based on control flow graph; data-flow analysis of the vulnerability detection algorithm can be effectively used to detect XSS, SQL injection vulnerabilities which exist in the source code. The blacklist is applied to check the input data in OK3W and Leichinews program. They have a common function to check all input string. The programs produce a certain false positive. Despite the weaknesses found in the report contains false positives, reporting the total number of articles, or less, from the relatively small number of reports of these find the true vulnerabilities have been greatly reduced the workload.

VI. CONCLUSIONS

Software needs to be secure in order to allow parties of different trust levels to interact with each other, without risking that un-trusted parties exploit critical parts of the software. Web applications are mostly susceptible to input validation vulnerabilities, also known as taint-style vulnerabilities, the most common of which are cross-site scripting and SQL injection. Because web applications are

made to be easily accessible through the internet, they are particularly exposed to attacks.

Static analysis allows programmers to look for security problems in their source code, without the need to execute it. Static analysis tools will always produce false positives and false negatives, because they need to make trade-offs between accuracy and speed, and because program analysis is inherently un-decidable.

This tool has manifests its usefulness in examining the web sites based on ASP of the virtual host computer in a high school. Despite the fact that fault rate remains high now, we would use data-flow analysis and add rules to detect sensitive information so as to yield higher accuracy of examination in source code and lower false positive amount within an acceptable bound.

REFERENCES

- [1] Brian Chess, Jacob West (2007). "Secure Programming with Static Analysis". Addison-Wesley. ISBN 978-0321424778.
- [2] Yao-Wen Huang, Shih-Kun Huang, Tsung-Po Lin, Chung-Hung Tsai. "Web Application Security Assessment by Fault Injection and Behavior Monitoring". In Proceedings of the Twelfth International Conference on World Wide Web (WWW2003), pages 148-159, May 21-25, Budapest, Hungary, 2003.

- [3] Yao-Wen Huang, Fang Yu, Christian Hang, Chung-Hung Tsai, D.T. Lee, Sy-Yen Kuo. Verifying Web Applications Using Bounded Model Checking". In Proceedings of the 2004 International Conference on Dependable Systems and Networks (DSN2004), pages 199-208, Florence, Italy, Jun 28-Jul 1, 2004.
- [4] Gray McGraw. Software Security. IEEE Security & Privacy[J]. March-April 2004,2(2):80-83
- [5] V. Haldar, D. Chandra, and M. Franz. Dynamic Taint Propagation for Java. In Twenty-First Annual Computer Security Applications Conference (ACSAC), 2005.
- [6] Yao-Wen Huang, Fang Yu, Christian Hang, Chung-Hung Tsai, D.T. Lee, Sy-Yen Kuo. Securing Web Application Code by Static Analysis and Runtime Protection". In Proceedings of the Thirteenth International World Wide Web Conference (WWW2004), pages 40-52, New York, May 17-22, 2004.
- [7] Steve Christey, Robert A. Martin. Vulnerability Type Distributions in CVE.[EB/OL].2007:V1.1, (2007 -5-22).
- [8] Viega, J., Bloch, J. T., Kohno, T. and McGraw, G. ITS4: A Static Vulnerability Scanner for C and C++ Code. In Proc. 16th Computer Security Applications Conferences pp. 257- 266, New Orleans, LA, 2000.
- [9] RATS: Rough Auditing Tool for Security. <http://www.securesoftware.com/resources/tools.html>.
- [10] Splint: Secure Programming Lint. <http://www.splint.org/>.
- [11] Wagner, D. BOON: Buffer Overrun Detection. <http://www.cs.berkeley.edu/~daw/boon/>.
- [12] FlawFinder Home Page. <http://www.dwheeler.com/flawfinder/>
- [13] OWASP: Top Ten Project. http://www.owasp.org/index.php/OWASP_Top_Ten
- [14] The WhiteHat Website Security Statistics Report - 8th Edition - Fall 2009. <http://www.whitehatsec.com/home/resource/stats.html>
- [15] David Scott, Richard Sharp. "Abstracting application-level web security". In Proceedings of the 11th international conference on World Wide Web (WWW2002), pages 396-407, Honolulu, Hawaii, May 17-22, 2002.
- [16] Xia Yiming. Security Vulnerability Detection Study Based on Static Analysis[J]. Computer Science, 2006, 33(10): 279-283.
- [17] Paul Biggar, David Gregg. "Static analysis of dynamic scripting languages". August, 2009

IMPLEMENTATION OF OBJECT ORIENTED APPROACH TO MODIFIED ANT ALGORITHM FOR TASK SCHEDULING IN GRID COMPUTING

CH SOMESWARA RAO¹, V MNSSV KR GUPTA², K.V.S. MURTHY³, J RAJANIKANTH⁴

¹⁻⁴Department of CSE, S.R.K.R Engg. College, Affiliated to Andhra University, Bhimavaram,
W.G.District, Pin-534 204, A.P. INDIA

Abstract— Now a day's Resource management and task scheduling are very important and complex problems in grid computing environment. It is necessary to predict resource state to get proper task scheduling. In this paper we propose object oriented modified ANT algorithm is a new heuristic algorithm. The inherent parallelism and scalability make the modified ANT algorithm very suitable to be used in grid computing task scheduling, whose structure is changes dynamically almost all the time. The scalability of proposed algorithm is validating by proposing a simple Grid simulation architecture for resource management and task scheduling. The algorithm when implemented in simulation environment produced good results in terms of minimal response time; maximum resource average utilization and task fulfill proportion.

Index Terms— Grid Computing, Task Scheduling, ANT Algorithm, Modified ANT Algorithm, Resource Management.

I. INTRODUCTION

Grids enable the sharing, selection, and aggregation of a wide variety of resources including super computers, storage systems, data sources, and specialized devices that are geographically distributed and owned by different organizations for solving large-scale computational and data intensive problems in science, engineering, and commerce. Grid computing, most simply stated, is distributed computing taken to the next evolutionary level [1,2].

Resource management and task scheduling are very important and complex problems in grid computing [3]. In grid environment, autonomy resources are linked through internet to form a huge virtual seamless environment. The resources in the grid are heterogeneous and the structure of the grid is changing almost all the time. In a grid, some resources may fail, some new resources enroll the grid, and some resources resume to work. So, it is necessary to do resource state prediction to get proper task scheduling. Since the static algorithms like round robin algorithm, fastest processor-largest task first algorithm etc, are no longer suitable for the effective utilization of the grid, a good task scheduling method should be used which is distributable, scalable and fault tolerant[4,5].

The aim of this paper is to simulate and analyze a task scheduling algorithm which is based on well known ant-algorithm. The task scheduling algorithm should work better in the grid environment, in which the structure of the grid changes dynamically almost all the time. It should reduce the total execution time of jobs submitted to the grid, by effectively scheduling the jobs on to the appropriate resources in the grid. It should also reduce the cost of using the resources to execute the jobs.

So, both the total execution time of the jobs and the cost of executing the jobs should be taken into consideration in scheduling the jobs on to the resources. If the cost of resources is not a factor modified forms of ant-algorithm, which can further reduce the total execution of the jobs.

A. ANT ALGORITHM

ANT algorithm is a new heuristic, predictive scheduling algorithm; it is based on the behavior of real ants. When the blind insects, such as ants look for food, every moving ant lays some pheromone on the path, then the pheromone on shorter path will be increased quickly, the quantity of pheromone on every path will affect the possibility of other ants to select path. At last all the ants will choose the shortest path.

- When a resource 'j' enrolls the Grid,

$$\tau_j(0) = m \times p + c / s_j$$

where 'm' is the No. of PEs, 'p' is the MIPS of one PE, 'c' is the size of the parameters, 's_j' is the parameter transfer time from resource 'j' to resource monitor.

- The pheromone on path from schedule centre to the corresponding resource will be changed as

$$\tau_j^{new} = \rho \cdot \tau_j^{old} + \Delta \tau_j$$

is the change of pheromone on path from the schedule center to resource j;

ρ is the permanence of pheromone ($0 < \rho < 1$)

$1 - \rho$ is evaporation of pheromone when task is assigned to resource j,

- = - k, k is the compute and transfer quality of the task.
- when task successfully returned from resource j,
- = Ce x k, Ce is the encourage factor.
- when task failed returned from resource j,
- = Cp x k, Cp is the punish factor.
- The possibility of task assignment to every resource will be recomputed as:

$$\rho_j^k(t) = \frac{[\tau_j(t)]^\alpha * [\eta_j(t)]^\beta}{\sum_u [\tau_u(t)]^\alpha * [\eta_u]^\beta}$$

$$= 0 \quad \text{others}$$

is the pheromone intensity on the path from schedule center to resource j

- $\eta_j(t)$ is the innate performance of the resource;
- α is the importance of the pheromone;
- β is the importance of resource innate attributes;

B. Drawbacks of the ANT algorithm

- The scheduler schedules a task based on the possibilities of the resources.
- The problem with this algorithm is, it may schedule a task to a resource with low possibility even if the resources with high possibility are free.
- But this randomness is required because if the tasks are always scheduled to the resource with high possibility, then the load on the resource may be increased and the jobs may be kept waiting in the queue for the resource to be free.
- It uses centralized scheduling scheme.

II. RELATED WORD

The efficient scheduling of independent computational jobs in a heterogeneous computing (HC) environment such as a computational grid is clearly important if good use is to be made of a valuable resource. Scheduling algorithms can be used in such a system for several different requirements [6]. The first, and most common, is for planning an efficient schedule for some set of jobs that are to be run at some time in the future, and to work out if sufficient time or computational resources are available to complete the run a priori. Static scheduling may also be useful for analysis of heterogeneous computing systems, to work out the effect that losing (or gaining) a particular piece of hardware, or some sub network of a grid for example, will have. Static scheduling techniques can also be used to evaluate the performance of a dynamic scheduling system after it has run, to check how effectively the system is using the resources available. The Ant Colony

Optimization (ACO) meta-heuristic was first described in[7] as a technique to solve the traveling salesman problem, and was inspired by the ability of real ant colonies to efficiently organize the foraging behavior of the colony using external chemical pheromone trails as a means of communication. ACO algorithms have since been widely employed on many other combinatorial optimization problems including several domains related to the problem in hand, such as bin packing and job shop scheduling, but ACO has not previously been applied to finding good job schedules in an HC environment.

Although various classification of task scheduling strategies exist, depending on the type of the grid, the scheduler organization and task scheduling strategy has to be chosen such that the resources in the grid are effectively utilized. Along with the scheduling organization, state estimation is also necessary for the dynamic grid. To achieve maximum resource utilization and high job throughput, re-scheduling of the jobs on different resources is also sometimes required.

Since the static algorithms like round robin algorithm, fastest processor-largest task first algorithm etc, are no longer suitable for the effective utilization of the grid, a good task scheduling method should be used which is distributable, scalable and fault tolerant.

In high throughput computing, the Grid is used to schedule large numbers of loosely coupled or independent jobs, with the goal of putting unused processor (resource) cycles to work. Build on the Internet and World Wide Web, the grid has emerged as a new class of infrastructure. By discovering and obtaining access to remote resources, the grid can provide scalable, secure, high performance calculating resources to make it possible for distributed scientific groups to work together to share resources in a large scale and to solve complex scientific problem that never practice in the previous time.

In grid, resources are distributed, heterogeneous, dynamic and instability. Dorigo et.al proposed [8], which is a new heuristic algorithm and based on the behavior of real ants. When the blind insects, such as ants look for food, the moving ant lays some pheromone on the ground, thus making the path it followed by a trail of this substance. While an isolated ant moves essentially at random, an ant encountering a previously laid trail can detect it and decide with high probability to follow it, thus reinforcing the trail with its own pheromone. The collective behavior that emerges means where the more the ants are following a trail, the more that trail becomes attractive for being followed. The process is thus characterized by a feedback loop, where the probability with which an ant chooses an optimum path

increases with the number of ants that chose the path in the preceding steps. These observations inspired a new type of algorithm called ant algorithms or ant systems.

The task-scheduling algorithm should work better in the grid environment, in which the structure of the grid changes dynamically almost all the time. It should reduce the total execution time of jobs submitted to the grid, by effectively scheduling the jobs on to the appropriate resources in the grid. It should also reduce the cost of using the resources to execute the jobs. So, both the total execution time of the jobs and the cost of executing the jobs should be taken into consideration in scheduling the jobs on to the resources. If the cost of resources is not a factor (situations like, all the resources of our interest belong to the same organization), modified forms of ant-algorithm, which can further reduce the total execution of the jobs.

III. SYSTEM ARCHITECTURE

In Grid environment, the client nodes can enroll the Grid at any time, deliver requests to the schedule center, and monitor the implementation of themselves tasks. There are five important modules in the schedule center; the architecture of the scheduler center can be expressed as shown in the Fig 1. Each node denotes a client or a Grid resource, and each edge denotes a link between nodes.

When a grid client delivers a request, the grid works as follows:

- The client delivers a request that contains an application description to the task receiver. The description is about

the work load of the application, communication load, and time limit, etc.

- The task receiver queues the tasks in priority, and delivers the first task in the queue to scheduling manager. The receiver maintains an unscheduled task queue; record the client name and user requirements, application workload, communication load, and time limit, etc.
- The scheduling manager selects a most appropriate scheduling scheme from all schemes according to the resource graph and user requirement, then delivers the scheme to task dispatcher and inform resource monitor. This is the most important and complex scheduling in the schedule center, there are many scheduling strategies can be put in the scheduling manager. We design the ant algorithm based strategy, and will discuss it in the following section.
- Task dispatcher delivers the task to the selected resource, and gives transfer delay and actual task assignment result to the resource monitor. The dispatcher maintains a scheduled task queue; record the assigned resource name, application work load, communication load, and time limit, etc. When some task is finished or failed, the dispatcher delete the task in the queue or put the task back to unscheduled task queue, and notifies the resource monitor.
- The resource monitor maintains the up to date of every resource and revises the resource graph.

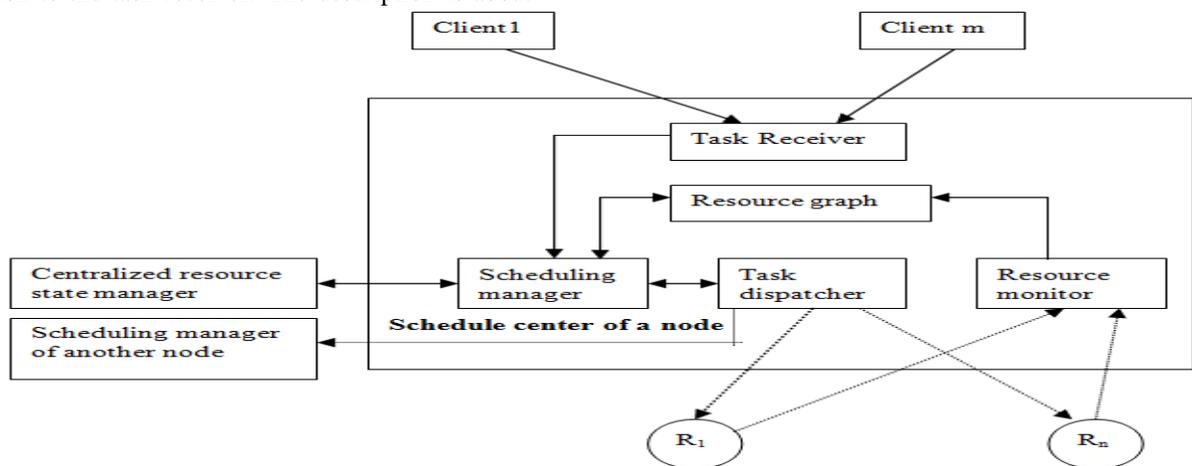


Fig 1 Detailed System Design

- We use a graphic to express the grid environment, each node denotes a client or a grid resource (number of processors, speed of each processor, I/O bandwidth, RAM capability, and disk capacity, etc.), and each edge denotes a link between nodes (bandwidth, delay, quantity of pheromone, etc.).

A. Improvements to the above said ANT algorithm

- The algorithm can be modified by using a threshold to minimize the total processing time.
- The processing costs of the tasks also can be controlled.

B. Modified Ant Algorithm

- When a resource ‘j’ enrolls the Grid,

$$\tau_j(0) = m \times p + c / s_j$$

where ‘m’ is the No. of PEs

‘p’ is the MIPS of one PE

‘c’ is the size of the parameters

‘s_j’ is the parameter transfer time from resource ‘j’ to resource monitor.

The pheromone on path from schedule centre to the corresponding resource will be changed as

$$\tau_j^{new} = \rho \cdot \tau_j^{old} + \Delta \tau_j$$

is the change of pheromone on path from the schedule center to resource j;

ρ is the permanence of pheromone ($0 < \rho < 1$)

$1 - \rho$ is evaporation of pheromone.

when task is assigned to resource j,

= - k, k is the compute and transfer quality of the task.

when task successfully returned from resource j,

= Ce x k, Ce is the encourage factor.

when task failed returned from resource j,

= Cp x k, Cp is the punish factor.

- The possibility of task assignment to every resource will be recomputed as:

$$\rho_j^k(t) = \frac{[\tau_j(t)]^\alpha * [\eta_j(t)]^\beta}{\sum_u [\tau_u(t)]^\alpha * [\eta_u]^\beta}$$

= 0 others

$\tau_j(t)$ is the pheromone intensity on the path from schedule center to resource j

$\eta_j(t)$ is the innate performance of the resource;

α is the importance of the pheromone;

β is the importance of resource innate attributes;

- The scheduler finds a resource ‘j’ on which a new task ‘k’ is to be scheduled at time ‘t’ with a

probability equal to it’s corresponding $\rho_j^k(t)$
until $(\rho_h^k(t) - \rho_j^k(t)) \leq T_h$.

where ‘h’ is the resource with highest $\rho^k(t)$.

T_h ’ is the threshold which will be taken as $1 / (\text{no. of resources})$.

- The scheduler finds a resource ‘i’ taking one resource at a time in the ascending order of $C_i / (\text{MIPS}_i)$ until $|\rho_j^k(t) - \rho_i^k(t)| \leq T_i * \ell$

where C_i is the cost of the resource ‘i’ per second
 MIPS_i is the total MIPS of the resource ‘i’

$T_i = T_h - (\rho_h^k(t) - \rho_j^k(t))$ and

‘ ℓ ’ is the cost reduction factor which is selected by the grid user who submits the task. $0 \leq \ell \leq 1$.

The scheduler schedules the new task ‘k’ on to the resource ‘i’.

IV. IMPLEMENTATION

In this paper we consider Grid Simulator(GridSim) toolkit, which supports modeling and simulation of a wide range of heterogeneous resources, such as single or multiprocessors, shared and distributed memory machines such as PCs, workstations, SMPs, and clusters managed by time or space-shared schedulers. GridSim can be used for modeling and simulation of application scheduling on various classes of parallel and distributed computing systems such as clusters, Grids, and P2P networks. The resources in clusters are located in a single administrative domain and managed by a single entity whereas, in Grid and P2P systems, resources are geographically distributed across multiple administrative domains with their own management policies and goals. The schedulers in cluster systems focus on enhancing overall system performance and utility, as they are responsible for the whole system.

The GridSim toolkit provides a comprehensive facility for simulation of different classes of heterogeneous resources, users, applications, resource brokers, and schedulers. It can be used to simulate application schedulers for single or multiple administrative domain(s) distributed computing systems such as clusters and Grids. Application schedulers in Grid environment, called resource brokers, perform resource discovery, selection, and aggregation of a diverse set of distributed resources for an individual user.

A. GridSim Entities

GridSim supports entities for simulation of single processor and multiprocessor, heterogeneous resources that can be configured as time or space shared systems. It allows setting their clock to

different time zones to simulate geographic distribution of resources. It supports entities that simulate networks used for communication among resources. The design and implementation issues of GridSim entities are discussed below:

i. User – Each instance of the User entity represents a Grid user. Each user may differ from the rest of the users with respect to the following characteristics:

- Types of job created e.g., job execution time, number of parametric replications, etc.,
- Scheduling optimization strategy e.g., minimization of cost, time, or both,
- Activity rate e.g., how often it creates new job,
- Time zone, and
- Absolute deadline and budget.

ii. Broker – Each user is connected to an instance of the Broker entity. Every job of a user is first submitted to its broker and the broker then schedules the parametric tasks according to the user’s scheduling policy. Before scheduling the tasks, the broker dynamically gets a list of available resources from the global directory entity. Every broker tries to optimize the policy of its user and therefore, brokers are expected to face extreme competition while gaining access to resources.

iii. Resource – Each instance of the Resource entity represents a Grid resource. Each resource may differ from the rest of resources with respect to the following characteristics:

- Number of processors;
- Cost of processing;
- Speed of processing;
- Internal process scheduling policy e.g., time shared or space shared;
- Local load factor; and
- Time zone.

iv. Grid Information Service – It provides resource registration services and maintains a list of resources available in the Grid.

v. Input and Output –The flow of information among the GridSim entities happen via their Input and Output entities. Every networked GridSim entity has I/O channels, which are used for establishing a link between the entity and its own Input and Output entities.

V. RESULTS

Graphical User Interface (GUI) is used for better usability of the system. In order to facilitate software (class) reuse, the GUI is solely implemented in a separate class. Another class is used for reading data (resource information) from user specified resource file. Resource information, user information, and task information should be maintained in the

Graphical user Interface. We give appropriate data for source info, resource info and user info, based on the given input we have apply the ANT, Round Rabin, and Modified ANT. Finally Fig 6 shows the Modified ANT algorithm is better in performance wise.

Fig 2. Source info

Fig 3 Resource Information

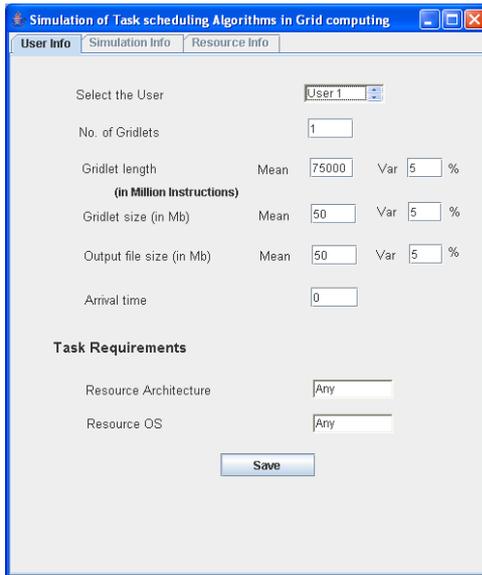


Fig 5 User information

| | Total Completion time | Total Cost |
|------------------------|-----------------------|------------|
| Ant Algorithm | 2,136.89 | 2,146 |
| Round-Robin | 3,213.12 | 4,277 |
| Modified Ant algorithm | 2,093.97 | 1,640 |

Fig 6. Comparison of different scheduling algorithms results:

VI. CONCLUSION AND FUTURE WORK

Since the structure of the grid dynamically changes, there is no particular scheduling algorithm, which can effectively utilize all the resources in a grid. But, the algorithm should be distributable, scalable and fault tolerant. It should also estimate the state of the resources, which are currently in the grid. And predictive state estimation is better than non-predictive state estimation because it uses the current state as well as the historical status of the resources. So, the static algorithms like round robin scheduling are no longer suitable.

The Ant algorithm is a heuristic predictive state estimating scheduling algorithm, which is distributable, scalable, and fault tolerant. The inherent parallelism and scalability make the algorithm very suitable to be used in grid computing task scheduling. The modified form of ant-algorithm can be used if cost of using the resources is not a concern. The selection of

threshold value in the modified ant-algorithm is very important.

As a future work, this scheduling method can be put into actual Grid environment for validation to make QOS scheduling. This algorithm can be made more suitable for wide use if pricing factor will also be included. The ant algorithm can also be applied to search for optimized path in computer networks.

References

- [1] Foster I, Kesselman C (eds.). The Grid: Blueprint for a Future Computing Infrastructure. Morgan Kaufmann: San Francisco, CA, 1999
- [2] Chetty M, Buyya R. Weaving computational Grids: How analogous are they with electrical Grids? Journal of Computing in Science and Engineering (CiSE) 2001; (July–August).
- [3] Sinha PK. Distributed Operating Systems: Concepts and Design. IEEE Press: New York, NY, 1997.
- [4] Ekemecic I, Tartalja I, Milutinovic V. A survey of heterogeneous computing: Concepts and systems. Proceedings of the IEEE 1996; **84**(8):1127–1144.
- [5] Buyya R, Abramson D, Giddy J. Nimrod/G: An architecture for a resource management and scheduling system in a global computational Grid. The 4th International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia'2000), Beijing, China, 2000. IEEE Computer Society Press: Los Alamitos, CA, 2000.
- [6] Braun TD, Siegel HJ, Beck N, Boloni LL, Maheswaran M, Reuther AI, Robertson JP, Theys MD, Yao B. A taxonomy for describing matching and scheduling heuristics for mixed-machine heterogeneous computing systems. Proceedings IEEE Workshop on Advances in Parallel and Distributed Systems, October 1998.
- [7] M. Dorigo et L.M. Gambardella, Ant Colony System : A Cooperative Learning Approach to the Traveling Salesman Problem, IEEE Transactions on Evolutionary Computation, volume 1, numéro 1, pages 53-66, 1997
- [8] M. Dorigo, V. Maniezzo, et A. Colomi, Ant system: optimization by a colony of cooperating agents, IEEE Transactions on Systems, Man, and Cybernetics--Part B , volume 26, numéro 1, pages 29-41, 1996.

Design of Compact U-Slot Circular Patch Antenna on RT DUROID 5880 Substrate

*¹K.V.L.Bhavani, ¹B.T.P.Madhav, ¹P.Poorna Priya, ¹V.V.S.Murty,
²K.Ravikumar, ³Srujana Adusumilli

¹Department of ECE, K.L.University, Guntur, A.P, India

²Department of ECM, K L University, Guntur, A.P, india

³Project Student Department of ECE, K L University, Guntur, A.P, india

Abstract-- This paper presents design and analysis of compact wire edge fed U-shape slot antenna. The proposed antenna has simple structure consisting U-shape slot on a circular patch of radius 13.1mm. The patch is designed on circular shape RTD substrate material of radius 13.2mm, with height of 5mm and whose permittivity is 2.2. By using only single patch a high impedance bandwidth is achieved. Simulated results show that the return loss is -16.15dB at the center frequency of exactly 3.282GHz and the simulated impedance bandwidth (VSWR<2) is 24%. The antenna is designed and simulated using CONCERTO software and theoretical results give good agreement with simulated results. Return loss, 2D, 3D gain, E and H Field distributions, pointing vector, polar plots, FD probe results, Electric and magnetic energy distribution, and quality factor are simulated for the proposed designed antenna was presented.

Keywords: U-slot micro strip patch antenna, circular patch, RTD substrate.

I. INTRODUCTION

Microstrip patch antenna in its simplest form consists of a radiating patch on one side of a dielectric substrate and a ground plane on the other side. Radiation from microstrip antennas occurs from the fringing fields between the edge of the microstrip antenna conductor and the ground plane [1]. Extensive research and development of microstrip antennas and arrays, exploiting the numerous advantages such as light weight, low volume, low cost, planar configuration, compatibility with integrated circuits have led to develop the proposed design model [2]. The proposed antenna design operates at 3.282GHz finds applications in the S-band which includes Microwave devices/communications, wireless

LAN, communication satellites, WLAN(Wi-Fi 802.11 a/n), wimax, Radars [3].

The disadvantage of MSPA is narrow bandwidth. To improve the bandwidth recent research efforts have been devoted. A high permittivity substrate could be used but this approach increases the coupling between the antenna and the ground plane. Embedding a suitable U-shaped slot in circular shape radiating patch is a very effective method for achieving a wide bandwidth. Several U-slot patch antennas have been reported recently to improve bandwidth [4-6]. In this paper U-slot is designed on circular patch to achieve more compactness and high impedance bandwidth.

Conventional Microstrip patch antenna designs with thick substrate layer causes major problem associated with impedance matching. The proposed antenna is designed on RT duroid 5880 substrate material which is low density, high weight material for high performance weight sensitive applications. The very low dielectric constant of RT/duroid 5880 laminates is uniform from panel to panel and is constant over a wide frequency range. Applications include airborne antenna system, light weight feed network, military radar systems, missile guidance system and point to point digital radio antennas [7].

The proposed antenna is designed using Concerto software. It is a state of the art system for high frequency field simulation. The main components are modeler, quickwave simulator, quickwave2D, CLASP, SOPRANO/EV and post processor. This provides a complete tool chain for RF and microwave electromagnetic design for use on 32 or 64 bit windows platform. Modeller is used to generate data and models for electromagnetic simulation [8].

II . DESIGN EQUATIONS

Fig 1 shows the circular patch MSPA.

A. Fields and currents: With no current sources the wave equation may be written as

$$(\nabla^2 + K^2)\vec{E} = 0 \text{ where } K = \omega\sqrt{\mu\epsilon}$$

Solution of the wave equation in cylindrical coordinates is

$$E_z = E_0 J_n(K\rho) \cos n\phi$$

Magnetic field components are

$$H_\rho = \frac{j}{\omega\mu} \frac{1}{\rho} \frac{\partial E_z}{\partial \phi} = -\frac{J_n}{\omega\mu\rho} E_0 J_n'(K\rho) \sin n\phi$$

$$H_\phi = \frac{-j}{\omega\mu} \frac{\partial E_z}{\partial \rho} = -\frac{J_n}{\omega\mu} E_0 J_n'(K\rho) \cos n\phi$$

$$E_\rho = E_\phi = H_z = 0$$

$\vec{K} = \hat{n} \times \vec{H} = \hat{\rho} H_\phi - \hat{\phi} H_\rho$ (Surface currents on the circular disk) at the edge of surface current $K_\rho(\rho=a) = H_\phi(\rho=a) = 0$

B. Resonant frequency

Resonant frequency of disk antenna for TM_{nm} mode may be evaluated from

$$f_r = \frac{K_{nm} C}{2\pi a_e \sqrt{\epsilon_r}}$$

Where K_{nm} the derivative of the Bessel function of order n & c is is velocity of light.

a_e is effective radius =

$$a \left[1 + \frac{2b}{\pi a \epsilon_r} \left(\ln \frac{\pi a}{2b} + 1.7726 \right) \right]^{\frac{1}{2}}$$

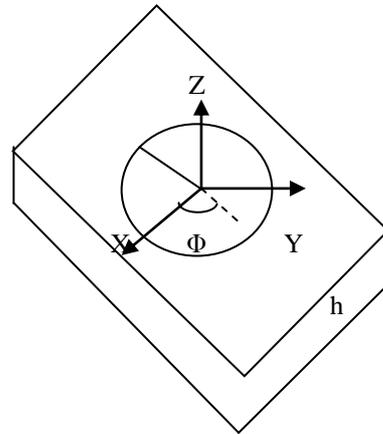


Fig 1. Circular MSPA

III ANTENNA DESIGN SPECIFICATION:

The geometrical configuration of the proposed U shape slot circular microstrip patch antenna is shown in Fig 2. Table 3.1 gives dimensions of the each element used to design the patch

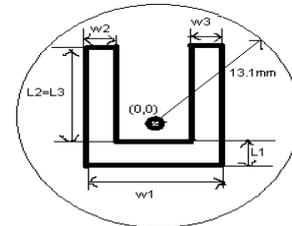


Fig 2: U shape rectangular microstrip patch

| S.No | Element | shape | Total Dimension (mm) | Xaxis (mm) | Yaxis (mm) |
|------|----------------|-------------|------------------------------|--------------|--------------|
| 1 | Patch | circular | Radius=13.1mm, height=0.05mm | 0 | 0 |
| 2 | slot1 (bottom) | rectangular | L1=4.6, w1=9 | -4.5 to 4.5 | -7.9 to -3.3 |
| 3 | slot2 (left) | rectangular | L2=6, w2=2.6 | -4.5 to -1.9 | -3.3 to -9.3 |
| 4 | Slot3 (right) | rectangular | L3=6, w3= 2.6 | 1.9 to 4.5 | -3.3 to 9.3 |
| 4 | Substrate | circular | Radius=13.2mm, height=5mm | 0 | 0 |
| 5 | Ground | circular | Radius=13.2mm, height=0.05mm | 0 | 0 |

| | | | | | |
|---|-----------|-----------|------------|---|---|
| 6 | Feed (dp) | Wire edge | Length=2,r | 0 | 0 |
| | | | adius=0.1 | | |

Table 3.1 U shape slot circular MSPA design values

Fig 2 shows U-shape slot circular patch designed on circular shape RT duroid substrate of radius 13.2mm whose permittivity value is 2.2 and 5mm height.

Fig 3a shows the top view of the designed patch model in CONCERTO and Fig 3b shows the side view of the designed patch.

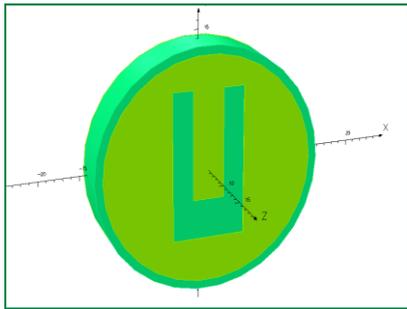


Fig 3a CONCERTO model for U shape slot circular microstrip patch antenna (top view)

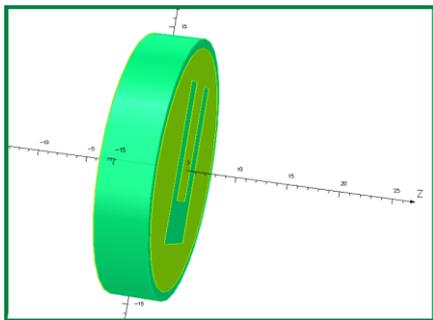


Fig 3b CONCERTO model for U shape slot circular microstrip patch antenna (side view)

The proposed antenna is fed at center (0,0). Wire edge feeding is used to reduce the complexity.

Fig 4 shows the side view of the proposed design with wire edge feeding.

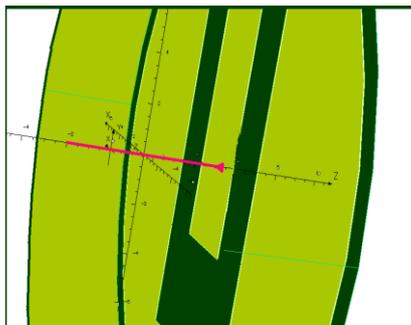


Fig 4: Feed location

IV. RESULTS AND DISCUSSION:

A. Return loss vs frequency

A good antenna might have a value of -10dB return loss as 90 % of the signal is absorbed and 10% is reflected back [10]. The proposed u-slot circular patch antenna is giving the excellent return loss curve in S-band. The curve has deep and wide dip at the center frequency of 3.282GHz. Fig 5 shows the return loss Vs frequency curve.

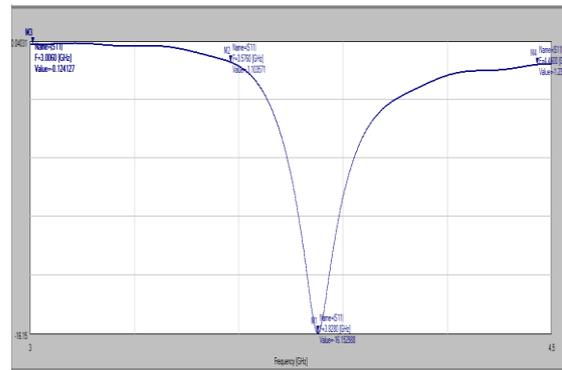


Fig 5: Return loss Vs frequency: -16.15dB at 3.282GHz.

B. 3D radiation pattern

Antennas will radiate the signal in one direction than others, so the gain is the amount of the power that can achieve in one direction at the expense of power lost in others [10]. Fig 6 shows 3D radiation pattern. Gain value obtained as 6.059dB

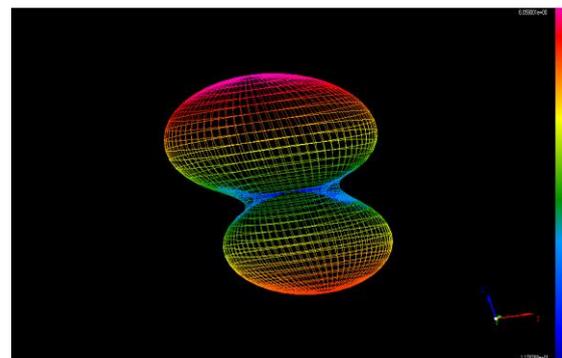


Fig 6: 3D radiation pattern: gain is 6.059dB

C. Radiation results in line form and polar form

Fig 7a and Fig 7b shows the radiation results in linear form when $\Phi=0^0$, $\theta = -90^0$ to 90^0 and when $\Phi=90^0$, $\theta = -90^0$ to 90^0 respectively.

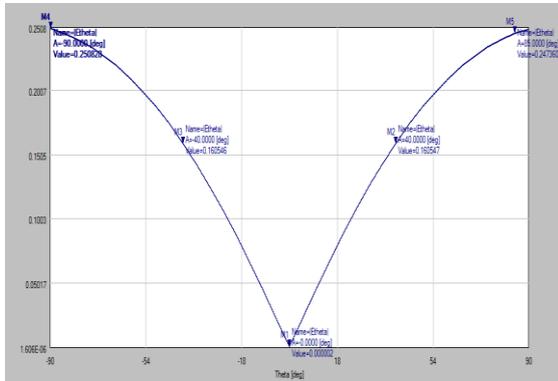


Fig 7a: Line form when $\Phi=0^0$ & $\theta =-90^0$ to 90^0

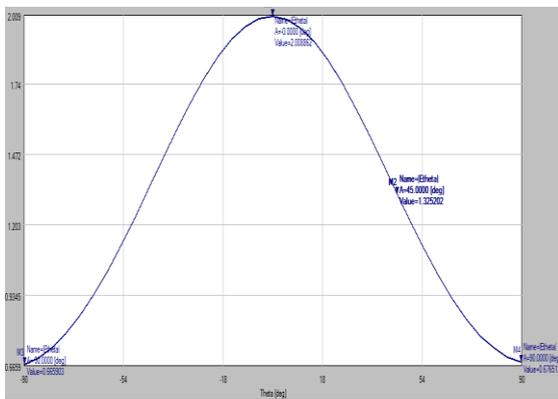


Fig 7b: Line form when $\Phi=90^0$ & $\theta =-90^0$ to 90^0

Fig 8a and Fig 8b shows the radiation results in polar form when $\Phi=0^0$, $\theta = -90^0$ to 90^0 and when $\Phi=90^0$, $\theta = -90^0$ to 90^0 respectively.

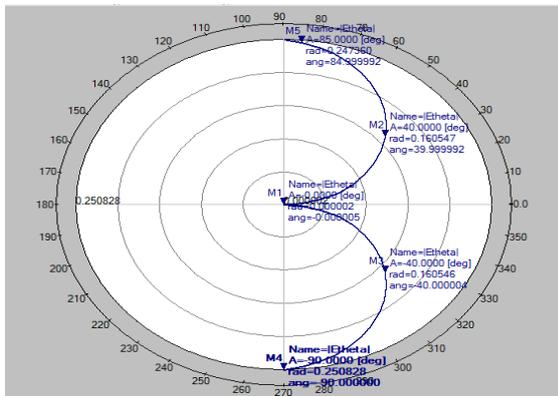


Fig 8a: polar results when $\Phi=0^0$ & $\theta =-90^0$ to 90^0

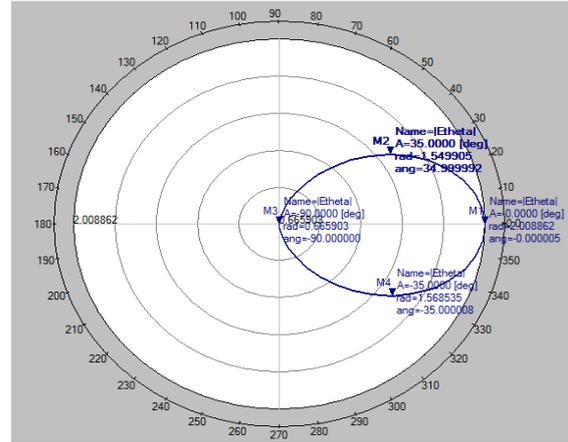


Fig 8b: polar results when $\Phi=90^0$ & $\theta =-90^0$ to 90^0

D. Field distributions and pointing vector:

Fig 9a, Fig 9b and Fig 9c shows the Electric field (E), magnetic field (H) and pointing vector $S=EXH$ respectively. It is observed that E-field distribution: $1.306327e^{-02}$ v/mm, H-field distribution: $1.937525e-02$ (A/mm) and Pointing vector: $3.754288e^{-06}$ W/mm²

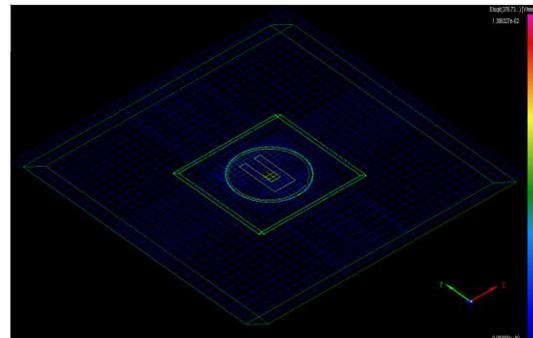


Fig 9a:E-field distribution : $1.30633e^{-02}$ v/mm

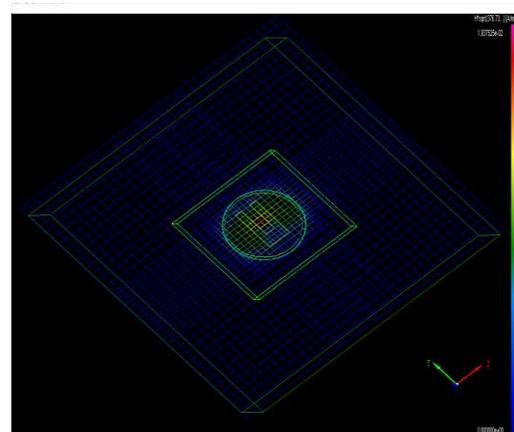


Fig 9b:H-field distribution : $1.9375e^{-02}$ (A/mm)

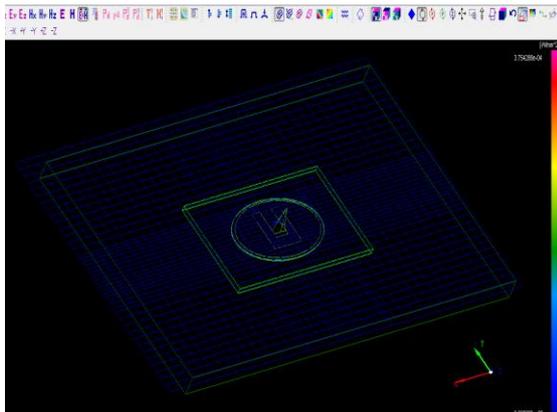


Fig 9c: Pointing vector: $3.754288e^{-06}$ W/mm²

E. Energy (Electric and magnetic) and quality factor: Fig 10 shows the energies and Q-factor

| File (J) | MIN | MAX | AIR | |
|------------|-----------------|-----------------|-----------------|---|
| Electric | 0 | 0 | 0 | E |
| Magnetic | 0 | 0 | 0 | M |
| Total | 0 | 0 | 0 | |
| Energy (J) | | | | |
| Electric | 4.491267e-005 | 4.491267e-005 | 4.491267e-005 | |
| Magnetic | 0.004391183 | 0.004391183 | 0.004391183 | |
| Total | 0.00440309 | 0.00440309 | 0.00440309 | T |
| Q-factor | | | | |
| | $1.2345678e-10$ | $1.2345678e-10$ | $1.2345678e-10$ | |

Fig 10: Energy and quality factor: Electric energy $4.491267e-005$ J Magnetic energy 0.004391183 J and total energy 0.00440309 J and quality factor is $1.2345678e-10$

F. FD probe results:

Fig 11a and Fig 11b shows that FD probe results in linear form and polar form respectively.

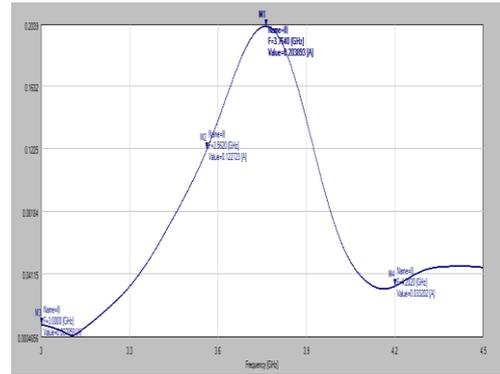


Fig 11a: FD probe results line form

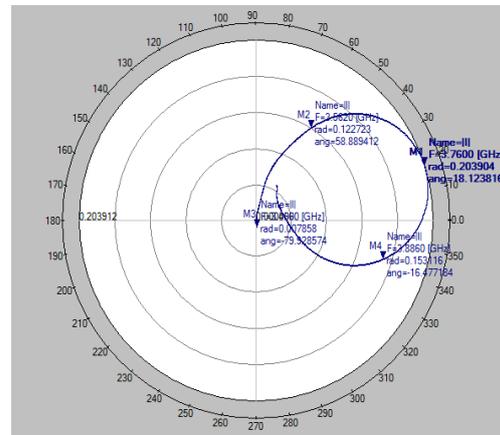


Fig 11b: FD probe results in polar form

V. CONCLUSIONS

In this paper the S-band U slot circular patch MSPA was designed by using RT duroid 5880 substrate material. From the simulated results, we can see that this proposed design giving good return loss of -16.15 dB at 3.282 GHz with high bandwidth. Main advantage of the proposed design is its compactness because of the circular patch shape so it can be used for all S-band applications includes wireless and Radar systems.

VI. ACKNOWLEDGEMENTS

The authors like to express their thanks to the management and department of ECE, K L University for their support and encouragement during this work.

VII. REFERENCES

[1]J.J.Bahl, P.Bhartia, "Microstrip Antennas Handbook", Artech House microwave library, 2000

[2] KIN-LU Wong, "Compact and broadband microstrip antennas", text book, John Wiley & Sons, Inc.

[3] *K.V.L.Bhavani, B.T.P.Madhav, P.Poorna priya, Prof. Habibulla Khan, G.Manoj Kumar, N.Durga Indira, "Gain Enhanced E- Shape Microstrip Patch Antenna at 2 GHZ on Air Substrate", International Journal of Advances in Science and Technology, ISSN 2229 5216 , Vol. 2, No.6, pp 51-58, June-2011.

[4] J A Ansari, Satya Kesh Dubey and Prabhakar Singh "Analysis of U-slot loaded Patch for Dual band Operation "International journal of microwave and optical technology vol. 3, no. 2, April 2008

[5] Wenwen Chail, Xiaojuan Zhang, and Jibang Liu "A Novel Wideband Antenna Design Using U-slot "PIERS online, vol. 3, no. 7, 2007

[6] RG Madhuri, PM Hadalgi, and SL Mallikarjun "U-Slot Rectangular Dielectric Resonator Antenna for Wideband Applications "International Journal of Electronics Engineering, 2 (2), 2010, pp. 249 – 252

[7] B.T.P. Madhav, V.G.K.M. Pisipati, K. Sarat Kumar, Habibulla Khan, "Comparative Study of Microstrip Rectangular Patch Array Antenna on Liquid Crystal Polymer and Rt-Duroid Substrates" International Journal of Electronics and Communication Engineering. ISSN 0974-2166 Volume 4, Number 2 (2011), pp.161-170

[8] www.concerto.com

[9] K.V.L.Bhavani, B.T.P.Madhav, P.Poorna Priya, Y. Joseph Manoj Reddy, N.Srinivas Sri Chaitanya, N.Krishna Chaitanya, ANALYSIS OF ORTHOGONAL FEED DUAL FREQUENCY RECTANGULAR MICROSTRIP PATCH ANTENNA FOR S-BAND APPLICATIONS, International Journal of Advances in Engineering Research <http://www.ijaer.com/> (IJAER) 2011, Vol. No. 1, Issue No. V, June ISSN: 2231-5152

[10] B.T.P.Madhav, VGKM Pisipati, K. Sarat Kumar, K.V.L.Bhavani, VGNS Prasad, "Log-periodic Toothed Planar Antenna on LCP for Ultra Wide Band Application" ,(IJAEST) International journal of advanced engineering sciences and technologies Vol No. 5, Issue No. 1, 062 – 066(2011) ISSN: 2230-7818

Author's Details:



K.V.L.Bhavani was born in India. A.P., in 1985. She received the B.Tech, M.Tech degrees from A.N.U., A.P., India in 2006, 2009 respectively. She is

pursuing her PhD in Microstrip Patch Antennas from K L University. From 2006 to till date she is working as Assistant professor in ECE branch K.L. University. She published Five International, one national Journal paper and five national conference papers. Her research interests include antennas and wireless communications.



B.T.P.Madhav was born in India, A.P, in 1981. He received the B.Sc, M.Sc, MBA, M.Tech degrees from Nagarjuna University, A.P, India in 2001, 2003, 2007, and 2009 respectively. From 2003-2007 he worked as lecturer and from 2007 to till date he is working as Assistant professor in Electronics Engineering. He has published more than 35 papers in International and National journals. His research interests include antennas, liquid crystals applications and wireless communications.



P.Poorna priya is working as Assistant professor in the ECE Department of KL University. Priya completed her M.Sc from Andhra University and M.Tech from KL college of Engineering. She is pursuing her PhD in Microstrip Slot Antennas from K L University. She published one International, one national Journal paper and two national conference papers She has been Involved in Teaching from last three years to UG students. Her field of Interest is in Communication systems, antennas and Image Processing.



V.V.S.Murthy was born on 02 January, 1981. He received his B.E. and M.Tech degrees in 2002 and 2006 respectively. He is a life member of IETE and ISTE. His research areas include Antennas and Radio wave propagation and optical image processing. Currently he is working as Associate Professor in ECE department of K.L. University, Guntur.

Performance of Content Based Mining Approach for Multi-lingual Textual Data

Kolla Bhanu Prakash

Research Scholar
Sathyabama University
Chennai, INDIA

M.A.Dorai Rangaswamy

Dept. of Computer Science and Eng.
AVIT
INDIA
Member IEEE

Arun Raja Raman

Retd. Professor, IITM
Chennai, INDIA
Member IEEE

Abstract— Data mining has become a necessary and powerful tool in the present era of web and internet communications. It has also evolved into media mining wherein heterogeneous data inputs like figures, videos and audios are gradually getting embedded into the web and this makes it quite complex and different. These and other aspects like currency and ‘liveliness’ of the web bring in more interesting features making a shift from translation to especially content extraction. Content extraction in web pages with Indian regional languages or English as the parent language have many aspects like free use of one language in another like ‘computer’ being used as it is with regional text and inclusion of other forms of data like hand written texts or sketches or drawings. This is common in education, news and entertainment and the focus of the current paper is in extracting content in a hybrid document with hand-written texts embedded. Work has been carried out initially with web documents in the form of computer generated text since they are more crisp in nature. Extending the idea, the present paper discusses on the results of hand-written text format and a comparative study with computer generated text format, which are less crisp in nature and more fuzzy depending on the writer. Beginning with letters having common content to words with common content, results of features on pixel maps are presented first. Later extraction using normalisation studies and classification means are presented.

I. INTRODUCTION

Web communication is becoming increasingly a powerful medium in variety of disciplines and with the spread of mobile and ad-hoc networks, it is an essential component in many areas of application. But with English as the main language used in the development in many of these conceptual and innovative applications, its adoption in regional and multi-lingual level needs more and more extensive work. One of the main problems here is in assessing the content of a web document and NOT the translated version of it, which may take more time when one searches content related information on-line perspective. So a content mining approach based on the file format of the web document is needed and this is developed so that the user or the node can react immediately for getting an in-depth view of a particular aspect in the document. In an earlier study, this idea was given for web documents having computer-generated texts. But many times the web documents may contain hand written texts in a different language or same language. It is the focus of the present study to look into aspects dealing with web documents

either in English or in a regional language like Tamizh or Telugu or Hindi, prepared in different modes.[1,2,3,4,5] The study and results are presented for letter content in a text with three variations.

II. CHARACTER OF PRESENT DAY WEB DOCUMENTS

Web documents are prepared in different ways with HTML occupying a standard form for developing web pages. But if one looks at the documents generated by a browser for presenting various aspects, the contents might differ. Fig.1 shows a typical multi-lingual web pages with varying characteristics. In Fig.1 (a) the web document shows image, icon and description in two languages in Tamizh and on the right hand side in English and it may be noted that left text is a translation of the right text in English. Fig.1 (b) shows another web page where the texts in regional language are literal replication of what is written in English like the word ‘computer’, being used in all language texts.



a) Multi-lingual web page with translation



b) Multi-lingual web page with replication. in regional text

Figure.1 Variations in multi-lingual Web pages

Even confining to one language like English, web pages in different regions show different content, depending on which is current in that region. This divergence is shown in “Fig.2”.The web page in World on top is completely different from the one for Asia, which again is different from what is in USA as shown in “Fig.2”.



Fig.2 Web pages on the same day in different regions.

So content extraction is more needed than literal translation of the document.[6,7,8] Many times the format of the web page is such that video, audio and text are in built in such a way that content is very apparent from the audio and video so that the user or node can decide immediately which is his need for further browsing. With this in view, a method based on pixel maps alone which any computer can ‘understand’ and can use to extract content, was developed and detailed elsewhere [17]. In this study the performance of content extraction with reference to text and character variations of different languages are discussed to form the basis for classification and training.

III. CONTENT BASED APPROACH FOR TEXTUAL DOCUMENT

Pixel maps of any text or figure or audio form the basis of storage transfer and interaction in any computer and though many formats are there beginning with .bmp to .png or .jpg, the jpg format has become universal for transfer, download and interaction between browsers and other application softwares. Here these are considered as the bases for content extraction. Beginning with letters and then on to words, content similarity is in different levels of usage and communication. So the study classifies text of different languages into three cases viz., a) letters having same content in most of the languages –CER(Content same in English and Regional language) , (b) letters unique to English and not present in others-CE and c) letters peculiar to other languages but not present in English. –CR. Later the study uses the pixel maps to convert into three parameter vectors defining attribute of that pixel map and later normalized with reference to parent one to get an index for the study and training[12,13,14].

Every pixel map after generating the code returns a vector consisting of three values. The first value indicates the occupancy ratio of pixels at the top portion, second value indicates the occupancy ratio for the center portion and the third value indicates the occupancy ratio of pixels at the bottom portion respectively. The sum of three values will be equivalent to 1.

In the present example pixel maps in four languages, English, Tamil, Telugu and Hindi, are taken to get four vectors resulting in a 3x4 matrix. Later these vectors are normalised with parent language say English and a matrix of order 4x4, is generated and this done by differential and ratio-based normalisation. In this matrix, the variations among the elements of the diagonal are of great use for predicting similarity in content and this is given in the present study.

IV. PERFORMANCE OF CONTENT BASED APPROACH

Letters and words in different languages have their own unique and distinctive features; but with English dominating the web in the last two decades, a tendency to use words mutually in English and regional languages has become popular. For example the word ‘computer’ is used as it is in many languages and communication. So content extraction calls for similar and dissimilar features in letters and words for better assessment of pixel map attributes. As mentioned before, the classification CER, CE and CR type interpretation is used for content study. Examples of CER are shown in Fig. 3 where letters have same content. Letters unique in English-CE- are shown in “Fig.4” and letters unique in other languages-CR- are shown in “Fig.5”



Fig.3. Pixel-maps of characters in four languages of same –CER-content



Figure 4. Letters unique to English-CE



Figure 5. Letters unique to regional language-CR

Normally text in any languages consists of words formed in a certain structured way and each of these words consists of characters native to the language in which the text is prepared. So it is preferable to look at extraction of features in characters[14,15,16,17] and here one can see how it is quite complicated between English and any regional language like Tamil or Telugu or Hindi.

Fig.6 gives a comparison of individual features of four letters in four languages having the same content in computer generated text format.

This gives us a clear idea of feature extraction in four different languages taken into consideration same content.

Since regional language letters have characters surrounding the main body, the pixel map is divided into three segments like 25% top, 50% middle and 25% bottom. Letters ‘g’ and ‘y’ in English have bottom 25% for example.

Later on we compare this with the handwritten text format which gave interesting features due its different crisp nature.

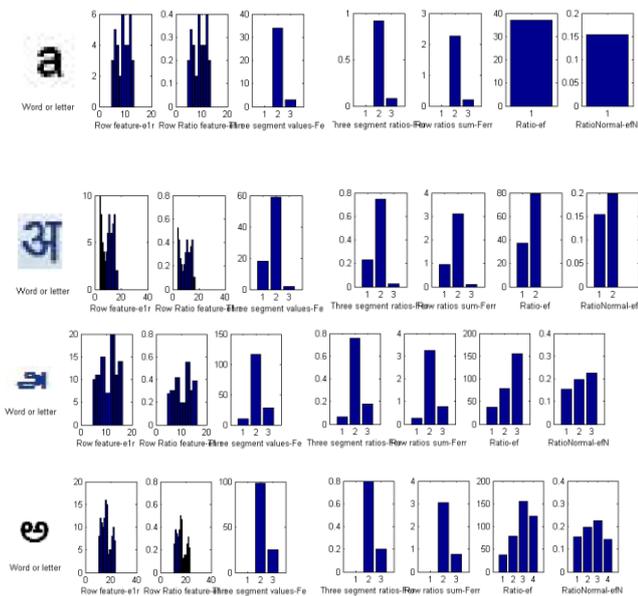


Figure 6. Feature extraction for CER-letter 'a' in CG format

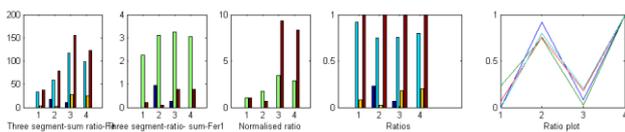
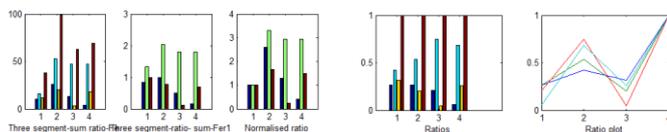
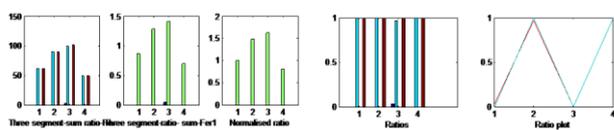


Fig.7 Feature comparison in letters in CER

Fig.7 clearly shows the variation in features in three segments and this can be used for classification and training. Similar figures could be obtained for CE and CR texts and these are shown in Fig.8 given below.



a) Feature comparison for CE letter 'x'



b) Feature comparison for CR-letter 'ke'

Fig.8 Feature comparison for CE, CR.

Now the performance of this approach is assessed through a new pixel map and the possibility of that belonging to CER, CE or CR is presented in terms on variations.

V.RESULTS AND DISCUSSION

Pixel map of letter 'y' is taken for assessing the performance as it is clear that it does not belong to any of the three

categories CER, CE and CR mentioned earlier. The vector representing the letter in terms the basic three parameters is taken and normalization with pixel maps used as parent ones is done and after converting to 3x3 matrix variations are presented for both difference and ratio approach. The variations are shown in Fig. 9 with CER CE and CR values.

| Type | Range(Dmatrix) | Range(Rmatrix) |
|----------|-----------------|-------------------|
| CER | -1.33% to 4.91% | 41.43% |
| CE | 0% to 18.07% | 41.27% |
| CR | -1.78% to 8.45% | 24.91% |
| 'y'(CER) | -3.1% to 5.36% | 0.2513 to 2.3328% |

Fig. 9 Performance of 'Y' pixel map

Once a new pixel map is later taken in four languages and the minimum and maximum variation is observed in the diagonal values of the resultant matrix, say for eg., letter 'x' for differences the minimum and maximum variation is found to be 0.0009 and 0.1679 respectively; whereas for letter 'a' the minimum and maximum variation is found to be 0 and 0.1769 respectively.

So, We have two ways of identification one by difference which is algebraic and another which is more rational. So the conclusion for 'x' is a) if we use diffMatrix attribute probability of 'x' belonging to CER is to the extent of $(.0009+.1679)/(0+.1769)$ which is 90% and this means content is similar to 'a'. b) if we use ratiMatrix the value is $7.1/.63 > 1$ so content is not 'a'. Similar approach can be applied to any new pixel map and we can find out whether the given pixel map belongs to CER or not.

In order to assess the performance of CER,CE and CR in the present paper we have taken a new pixel map 'y' and did the same procedure and made a comparison with all the three cases in both difference and ratio. The detailed study is given below.

In the case of differences for CER the variation is found to be 63%, for CE the variation is 52%, for CR the variation is 19%. So, in this category we can give a conclusion that the considered pixel map may belong to the category CER.

In the case of ratios for CER and CE the variation is found to be 41% and for CR the variation is 25%. So, the probability that the considered pixel map may belong to CER is more compared to the other two categories. The variations are shown in the form of histograms in Figures.10(a) and 10(b).

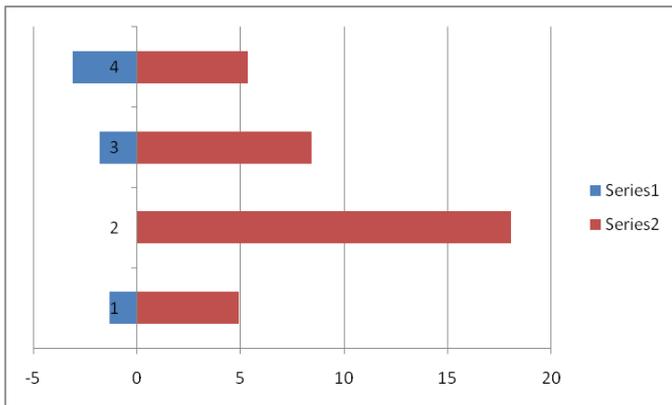


Fig.10(a) Feature comparison for differences

In observing both the categories differences and ratios we can conclude strongly that the pixel map 'y' may belong to CER category, which says about the content unique to English and other regional languages.

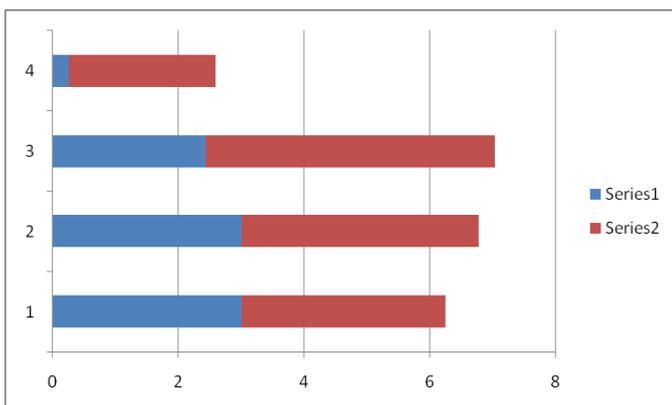


Fig.10(b) Feature comparison for ratios

Similar approach can be applied to any new pixel map and we can find out whether the given pixel map belongs to CER, CE or CR.

In the next case we combined all CER, CE and CR into one category which tells us about single concept that is all of them talk about same content. Same pixel map 'y' is taken for analysis of performance. And it is observed that for differences the probability is 40% and for ratios the probability that 'y' belongs to the above category is 37%. Since, the probability is very less we can say that 'y' may not belong to the above category.

Since the web document can contain hand written, computer generated or both the formats, the model that is discussed here should support any of the formats. Identifying the features obtained in both formats and comparing the both it is observed that 10% variation at the top portion, less than 5% variation in the middle portion and 20% variation in the bottom portion between both the formats for CER letter 'a'.

So, in the present discussion we are converting hand written text features into computer generated format. This can be done in two different ways. First ratios between the both

formats can be calculated. Second differences between both the formats are calculated. Since the variations may not be completely satisfactory, these can be further normalized to give better results.

Till now all the discussion given above is with computer generated texts. But, often we find hand written text also available as web documents. So, as a final study we compared the pixel map with the handwritten text category also and the conclusion for this is given below.

It is observed that for differences the probability is 43% and for ratios the probability that 'y' belongs to the above category is 40%. Since, the probability is very less we can say that 'y' may not belong to the above category.

VI. CONCLUSIONS

Extraction of content in multi-lingual web documents is essential for education and other activities on the net so that the user can surf on interested areas immediately. A method based on feature extraction for words in multi-lingual documents is developed and the complexities and numerical aspects are discussed for typical examples. The examples are from letters to words bringing out the need to include character variations in developing the mining approach.

REFERENCES

- [1] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins "Digital image processing using matlab", 2002.
- [2] Renu dhir "Feature extraction and classification for bilingual script (Gurumukhi and Roman)", April 2007.
- [3] Bing Zhao, Stephen Vogel "Adaptive parallel sentences mining from web bilingual news collection", 2002.
- [4] S.-C. Chen, S. H. Rubin, M.-L. Shyu and C. Zhang, A dynamic user concept pattern learning framework for content-based image retrieval, IEEE Transactions on Systems, Man, and Cybernetics: Part C 36(6) 2006) 772-783.
- [5] Z.-N. Li and M. S. Drew, Fundamentals of Multimedia (Prentice Hall, NJ, 2004).
- [6] L. Pan and C. N. Zhang, A criterion-based role-based multilayer access control model for multimedia applications, in Proc. Eighth IEEE Int. Symposium on Multimedia (San Diego, CA, USA, 2006), pp. 145-152.
- [7] G. Lu, Multimedia Database Management Systems (Artech House Publishers, Boston/London, 1999).
- [8] Y. Li, C.-C. J. Kuo and X. Wan, Introduction to content-based image retrieval —Overview of key techniques, in Image Databases: Search and Retrieval of Digital Imagery, eds. V. Castelli and L. D. Bergman (John Wiley, New York, 2002), pp. 261-284.
- [9] Q. Iqbal and J. K. Aggarwal, CIRES: A system for content-based retrieval in digital image libraries, in Proc. Int. Conf. Control, Automation, Robotics and Vision (ICARCV) (Singapore, 2002), pp. 205-210.
- [10] A. Kuchinsky, C. Pering, M. Creech, D. Freeze, B. Serra J. Gwizdka, Fotofile: A consumer multimedia organization and Retrieval system, in Proc. ACM CHI Conference (New York, NY, USA, 1999), pp. 496-503.
- [11] A. Gupta and R. Jain, Visual information retrieval, Communications the ACM 40(5) (1997) 71-79.
- [12] A. Pentland, R. W. Picard and A. Sclaroff, Photobook: Content based manipulation of image databases, Int. J. Computer Vision 18(3) (1996) 233-254.

- [13] Kolla Bhanu Prakash, M.A.Dorai Ranga Swamy, Arun Raja Raman
“Mining approach for documents containing multilingual Indian
texts”(NCRTAC-09), Bharath University, Chennai.
- [14] Kolla Bhanu Prakash, M.A.Dorai Ranga Swamy, Arun Raja Raman “A
Two-Input Neuron model for documents containing multilingual
Indiantexts” (EPPCSIT-09), Guru Nanak Dev Engineering College,
Ludhiana.
- [15] Kolla Bhanu Prakash, M.A.Dorai Ranga Swamy, Arun Raja Raman
“Feature extraction for content mining in multi-lingual documents”
(NCICN 2010), Sathyabama University, Chennai.
- [16] Kolla Bhanu Prakash, M.A.Dorai Ranga Swamy, Arun Raja Raman
“Text Studies Towards Multi-lingual Content Mining for Web
Communication” (TISC2010), Sathyabama University, Chennai.
- [17] Kolla Bhanu Prakash, M.A.Dorai Ranga Swamy, Arun Raja Raman
“Content Extraction for Multi-lingual Web documents” CIT Journal of
Research, Volume 1, Issue 3, NOV 2010, pp.93-101.

A Robust Kalman Filter Based Sensorless Vector Control of PMSM with LMS Fuzzy Technique

Bindu V¹, A Unnikrishnan², R Gopikakumari¹

¹ Division of Electronics, School of Engineering, Cochin University of science & Technology, Kochi, India. 682022

² Naval Physical and Oceanographic Laboratory, Kochi, India .682021

Abstract—In this paper a fuzzy logic technique optimized with Least mean square (LMS) algorithm is used for the vector control of Permanent magnet synchronous motor (PMSM). It is made sensorless by the use of Extended Kalman filter (EKF) algorithm which estimates the speed, rotor position, direct and quadrature axis current of PMSM. Unlike in the previous approaches the present work uses only the sensed line currents as measurements, and thus following a blind system identification approach. The Least mean square algorithm incorporated along with fuzzy inference, optimizes the weights used for combining the rules, which in turn makes the controller more efficient. The results show the improvement in control algorithm when LMS technique is incorporated with fuzzy decision process. The Lyapunov exponent calculated on the phase plane generated by direct axis current, quadrature axis current and motor speed, shows that the system is always asymptotically stable.

Key words — blind system identification, Extended Kalman filter, Fuzzy logic speed control, least Mean Square algorithm, Lyapunov exponent.

1. Introduction

High torque to inertia ratio, superior power density, high efficiency and many other advantages made PMSM the most widely acceptable electrical motor in industrial applications. The invention of Vector control made the ac drives equivalent to DC drives in the independent control of flux and torque. To facilitate vector control the stator quantities are resolved into components which rotate in synchronism with the rotor. For this transformation of stator quantities into synchronously rotating frame, the accurate knowledge of speed and rotor position is required. It usually requires mechanical sensors for measurement of speed in variable speed applications. But these types of shaft mounted mechanical sensors will make the system more complex and moreover reduces the reliability of the drive system. Accordingly, sensor less operation of PMSM has been receiving wide attention recently [1] in variable speed drives.

In this paper, we propose to estimate the motor quantities like speed, flux vector position, currents in direct and quadrature axes, from the measurements of three phase stator currents using an Extended Kalman filter (EKF). The proposed EKF estimation does not

require the information of the input voltage at all and it makes it different from similar other works[2]. So the present estimation treats the whole problem as blind system identification. The main problem associated with the EKF is its dependence on the parameters like the initial state, initial state co- variance, the measurement noise and the plant noise. The convergence is highly dependent on the choice of the covariance matrices that appear in EKF algorithm. Here the measurements to the EKF are the output of space vector PWM inverter, which is non sinusoidal and rich in harmonics, there by degrading the performance of EKF. To overcome this problem, in this paper a Bayesian approach is used in modifying the elements of the measurement covariance matrix, on the run [4]. The results show the robustness of the resulting implementation of the EKF.

Fuzzy control is different from the traditional PI control in the sense that it does not depend on precise system mathematical model [4]. The fuzzy logic technique has become very popular in the control of ac drives because of the flexibility in accommodating overlapping information in the definition of terms. Standard algorithm computes a fuzzy function on the basis of the error and change in error of the set speed and the estimated speed using a set of rules. The usual approach is to compute a fuzzy function on the error between the set speed and the estimated speed using a set of rules. In the usual approach all the rules are fired with equal weightings. As an optimization to FLC following [6], in this paper a gradient descent method is used, to adjust the weighting of each of the rules of the fuzzy controller, which minimizes the square error between the rotor speed and the reference speed. Least mean square based adaptive fuzzy along with Robust EKF is demonstrated in an environment, where the PMSM runs in Simulink and rest of the speed estimation and the control run concurrently in Matlab. And the results are compared with conventional Fuzzy logic used with conventional Kalman filter to show the improvement in performance. The stability of the controller is confirmed by calculating the Lyapunov exponent value.

2. Mathematical model of PMSM

The dynamic model developed on a synchronously rotating reference frame describes better the behavior of the motor for the vector control. Therefore the stator variables are transformed into a synchronously rotating d-q frame. The stator of the PMSM is similar to that of the wound rotor synchronous motor. The back emf produced by a permanent magnet is similar to that produced by an excited coil. A PMSM can be mathematically represented by the following equation in the d-q axis synchronously rotating rotor reference frame for assumed sinusoidal stator excitation [7]:

$$\frac{d}{dt} \begin{bmatrix} i_d \\ i_q \end{bmatrix} = \begin{bmatrix} R + pL_q & \\ -p\omega_r & L_q \end{bmatrix} \begin{bmatrix} i_d \\ i_q \end{bmatrix} + \begin{bmatrix} v_d \\ v_q \end{bmatrix} \quad (1)$$

$$d\theta_e/dt = P \quad (2)$$

$$T_e = 3P/2 [i_d i_q] \quad (3)$$

$$T_e = p \omega_r + T_l \quad (4)$$

Where v_d and v_q are the d, q axis voltages, L_d and L_q are the d,q axis inductances and i_d and i_q are the d,q axis stator currents, respectively. The other parameters are:

- R : the stator resistance per phase
- λ_m : the constant flux linkage due to rotor permanent magnet,
- ω_r : the angular rotor speed,
- θ_e : the rotor position in electrical degrees,
- P : the number of pole pairs of the motor,
- p : the differential operator,
- T_e : the developed electric torque,
- T_l : the load torque,
- σ : the rotor damping coefficient,
- J : the inertia constant

The current control is made possible through a vector control approach. In order to make the PMSM system linear, the d axis current is set to zero. So control of PMSM will become as easy as that of a DC motor. The d-q axis currents are related to the three phase stator currents by the equation:

$$\begin{bmatrix} i_d \\ i_q \end{bmatrix} = \begin{bmatrix} -\sin \theta_e & \cos \theta_e \\ \cos \theta_e & \sin \theta_e \end{bmatrix} \begin{bmatrix} 0 & 1/\sqrt{3} & -1/\sqrt{3} \\ 2/3 & -1/3 & -1/3 \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \quad (5)$$

3. Estimation of speed and rotor position using robust extended Kalman filter

The Extended Kalman filter [8] is an optimal recursive algorithm suitable to estimate the state of nonlinear dynamic systems. The system is described by the following state equations:

$$\dot{x} = f(x) + Bu_k \quad (6)$$

$$y = Cx + v_k \quad (7)$$

where v_k and w_k are the zero mean white Gaussian noise. And the state vector, $x = [i_d \ i_q \ \omega_r \ \theta_e]$. The measurements are the three phase stator currents $[i_a \ i_b \ i_c]$. And u is the input voltages v_d and v_q . Here the state vector x is augmented with the voltage inputs v_d, v_q . Then the new state vector X becomes $[i_d \ i_q \ \omega_e \ \theta_e \ v_d \ v_q]$. In discrete form the augmented state model is represented as:

$$\begin{bmatrix} i_d(k) \\ i_q(k) \\ \omega_r(k) \\ \theta_e(k) \\ v_d(k) \\ v_q(k) \end{bmatrix} = \begin{bmatrix} 1 - \frac{R}{L T_s} & \omega_r T_s & 0 & 0 & T_s/L & 0 \\ -\omega_r T_s & 1 - \frac{R}{L T_s} & (-\phi_f/L) T_s & 0 & 0 & T_s/L \\ 0 & (\frac{3\phi_f}{2J}) P^2 T_s & (-\frac{B}{J}) T_s & 0 & 0 & 0 \\ 0 & 0 & T_s & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i_d(k-1) \\ i_q(k-1) \\ \omega_r(k-1) \\ \theta_e(k-1) \\ v_d(k-1) \\ v_q(k-1) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_2(k) \\ w_3(k) \\ w_4(k) \\ w_5(k) \\ w_6(k) \end{bmatrix} \quad (8)$$

The mathematical model of PMSM is mutually coupled and hence nonlinear [7]. In the present work, an Extended Kalman filter is used for estimating the speed and the rotor position, from the non linear system given in (8), based on the measured values of line currents. For a given sampling time T_s , both the state estimate $\hat{x}_{k/k}$ and its covariance matrix $P_{k/k}$ are generated by the filter through a two step loop predictor corrector process.

The corrector algorithm starts with an initial value of \hat{x}_0 and follows as below

1. Computation of the Kalman gain

$$K_k = P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R)^{-1}$$
 where K_k is the Kalman gain for k^{th} iteration and $R = X =$ and R is a constant measurement noise covariance matrix.
2. Update estimate with the measurements

3. Updating the error covariance as

$$P_{k/k} = (I - H_k K_k) P_{k/k-1}$$
 where I is a identity matrix.

The predictor algorithm involves

1. Projecting the state error covariance matrix ahead for the next iteration as

$$P_{k/k+1} = A_k P_{k/k} A_k^T + Q$$

where $X = \begin{bmatrix} \omega \\ \theta \end{bmatrix}$ and Q is a constant process covariance matrix.

2. And the state is projected ahead as

$$X_{k+1} = f(X_k)$$

Beginning the iteration with an initial value of x is $X(0)$ and the Covariance $P(0)$ the Kalman filter estimates the values of flux angle θ and the speed ω . The convergence of Kalman filter is highly effected by the choice of $X(0)$, Q and R . Usually these matrices are chosen by trial and error approach. Beginning the iteration with an initial value of x is $X(0)$ and the Covariance $P(0)$ the Kalman filter estimates the values of flux angle θ and the speed ω . Since the measurements contain a large number of harmonics the performance of the EKF is often tainted. In this paper, the EKF is made more robust using an alternate method, based on Bayesian approach [4]. A scalar weight G_k is introduced for each data sample such that the variance of R is weighted with G_k , given by:

$$G_k = \frac{a}{b + (Z_k - \hat{h}_k \hat{x}_k/k)}$$
 (9)

a and b are taken as constants equal to one. This weight G_k is utilized to scale down the measurement covariance matrix R , before it goes to EKF algorithm for the calculation of Kalman gain i.e. R is modified as R/G_k .

During simulations, it has been observed that the robust EKF eliminates the need for manual parameter tuning of measurement covariance matrix in EKF equation. The results also show that the Kalman filter has become robust against input variation, with this technique.

4. LMS fuzzy logic speed controller

The over control of the speed is realized in terms of e and Δe . The error in speed and the rate of change of

speed error are considered as the input linguistic variables and the quadrature axis current is considered as the output linguistic variable. The support for all the fuzzy variables viz. e and Δe , variable of output ie i_q is scaled to $[-1, 1]$. In this case 7 membership functions are used, viz., NB, NM, NS, ZO, PS, PM, PB as shown in Fig.1

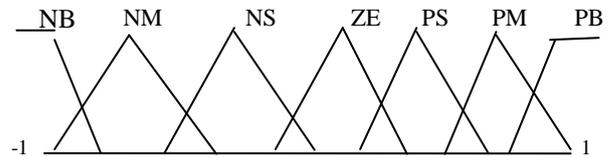


Figure. 1 Membership function for computing quadrature axis current.

Rules for computing the quadrature axis current is given below

TABLE 1
Rules of quadrature axis current

| i_q | e | | | | | | |
|------------|----|----|----|----|----|----|----|
| Δe | NB | NM | NS | ZO | PS | PM | PB |
| NB | NB | NB | NB | NB | NM | NS | ZO |
| NM | NB | NB | NB | NM | NS | ZO | PS |
| NS | NB | NB | NM | NS | ZO | PS | PB |
| ZO | NB | NM | NS | ZO | PS | PM | PB |
| PS | NM | NS | ZO | PS | PM | PB | PB |
| PM | NS | ZO | PS | PM | PB | PB | PB |
| PB | ZO | PS | PM | PB | PB | PB | PB |

The block diagram representation of the controller is shown in Fig 2.

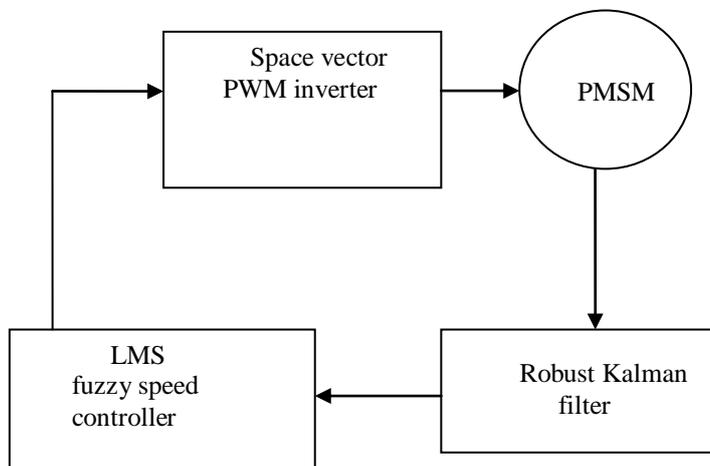


Figure 2 Block diagram representation of the control system

Using the product inference rule and central average defuzzifier method the fuzzy output is expressed as:

$$q = \frac{\sum_{i=1}^n \mu^i C_m}{\sum_{i=1}^n \mu^i} \quad (10)$$

Where n is the no. of rules, μ^i is the weighting factor which is to be adjusted and C_m is the membership function of i^{th} rule. The weighting parameters C of each rule are adjusted in order to minimize the square of the error between the instantaneous rotor speed and reference speed. For this the gradient descent method is used. The instantaneous error function is defined as:

$$J(k+1) = \frac{1}{2} [\omega_m(k+1)]^2 \quad (11)$$

where ω_m is the reference motor speed and the parameters of C_j are adjusted with [6]:

$$\Delta C_j(k) \approx \alpha Sgn(B) \quad (12)$$

where α is the learning rate. B is the motor parameter,

$$B = \dots ; \text{ and}$$

$$A = \exp(\dots)$$

where T is the sampling period. K_i and K_u are the scaling factors. The current command i_q^* is obtained by the output of AFC as:

$$i_q^*(k) = i_q^*(k-1) + k_u q(k) \quad (13)$$

5. Computation of the Lyapunov exponent

The Lyapunov exponent is an important indicator of the stability of a nonlinear system. Given a continuous dynamical system in an n dimensional phase space, we monitor the long term evolution of an infinitesimal sphere of initial condition. The i th one dimensional Lyapunov exponent is then defined as follows:

$$\lambda_i = \dots \quad (14)$$

The signs of the Lyapunov exponents provide a qualitative picture of a system's dynamics.

- If $\lambda_i = 0$; a marginally stable orbit.
- If $\lambda_i < 0$; a periodic orbit or a fixed point
- If $\lambda_i > 0$; Chaos

6. Simulation and Results

Simulation of the given PMSM has been carried out using Simulink. The three phase stator currents are taken from the motor model, and given to the robust extended Kalman filter. The Robust EKF estimates the instantaneous motor speed, rotor position, quadrature and direct axis currents and voltages. The estimated motor speed is compared with the reference speed and the error produced and change in the error is given to the Adaptive fuzzy speed controller. Where, it undergoes

the fuzzy inference process according to the rule generated in TABLE1. It gives quadrature axis current as the output linguistic variable. Fuzzy speed controller output is compared with the estimated quadrature axis current and the error produced is passed through a PI current to voltage converter to produce quadrature axis voltage. The direct axis current reference is set to zero and this value is compared with the estimated direct axis current and the error produced is passed through another PI current to voltage converter which produces direct axis voltage. The direct and quadrature axis voltages are converted into two axis stator voltages using the inverse park transformation. The two axis stator voltages are utilized to trigger the space vector PWM inverter. The space vector PWM inverter drives the PMSM. The approach towards the estimation of speed and rotor position is stable and converges very fast. It needs only the three phase stator current values which are easily available from the machine.

In order to check the stability of the system Lyapunov exponent was calculated and found that the system is always asymptotically stable.

The Lyapunov exponent values for the three phase vectors direct axis current, quadrature axis current, motor speed are as follows: -8.2999, -6.9554, -5.2624.

The phase plane trajectory of these three vectors is plotted in Figure 7.

The Robust EKF algorithm is developed in Matlab and integrated with Simulink using embedded Matlab facility. The parameters used for simulation are discussed below. The PMSM parameter used in this paper is **1.1 KW, 4 poles, R=2.875Ω, L=0.423H, =1.7wb/m², J= 0.008Kg-m²**. During the simulation, it was seen that convergence for the speed estimating Kalman filters is highly dependent on the initial values viz. $X(0)$ and $P(0^-)$. The values of those matrices are given as

$$X(0) = [0;0;0;0;0], R = \text{dia} [10; 10; 10]$$

$$P(0^-) = \text{dia} [200; 200; 200; 200; 200; 200]$$

The sampling time chosen is $4.5e^{-5}$ S. The values of proportional and integral gain constants used are, $K_p = 60$, $K_i = 2$; for inner quadrature axis current controller. $K_p = 40$, $K_i = 2$; for inner direct axis current controller.

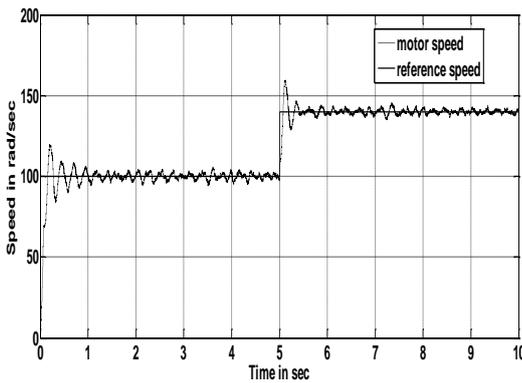


Figure 3. Step response of speed with LMS estimation of C_m in the fuzzy control strategy

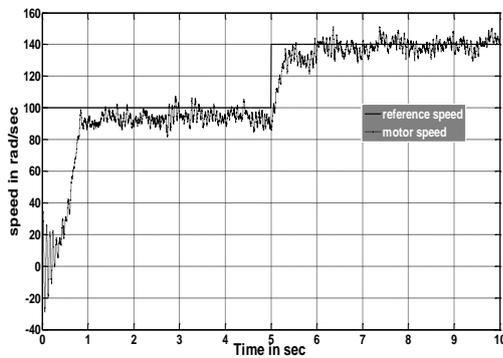


Figure 4. Step speed response with fuzzy control, without LMS estimation of C_m

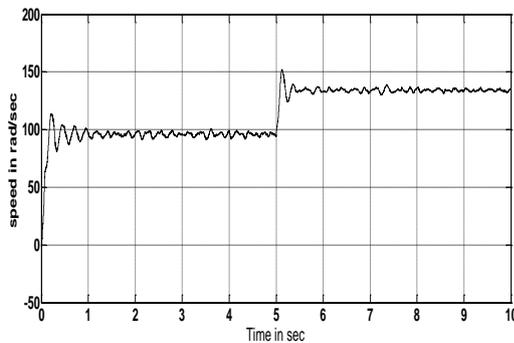


Figure 5. Estimated speed with Robust EKF

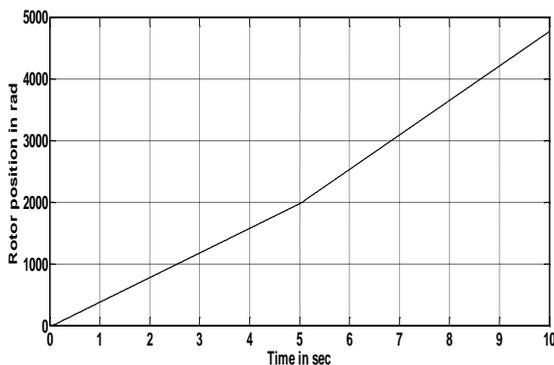


Figure6. Estimated rotor position with Robust EKF

Fig 3 shows the step response of mechanical speed of PMSM with a LMS Fuzzy, Robust EKF, from the result it is clear that the speed converges within 0.5 sec time with an error of 1%.hence it is very quick in operation compared to the conventional fuzzy technique. Fig 4 shows the same step response with conventional fuzzy, there it takes nearly 1 sec time to converge. Fig 5 shows the estimated motor speed for the step speed change of 100 rad/sec to 140 rad/sec with Robust EKF. From the results it is clear that the LMS technique incorporated along with Fuzzy logic made it quicker compared to the conventional case. In the speed graph it is clear that the estimated speed is completely coincident with the running speed of motor. Figure 6 shows that the slope of the estimated rotor position also changes when the speed steps into a new value. These step changes are required in electric vehicle application of the PMSM. The information regarding rotor position is utilized in Park inverse transformation. Fig 7 shows that the system is asymptotically stable.

The Robust EKF completely eliminates the difficulties in convergence of EKF algorithm. One of the main distinction from similar other works in literature is that the proposed method uses only three stator current measurements of PMSM in order to estimate six quantities in the state vector X , at a time. In order to assess the stability of the proposed control strategy, the phase plane corresponding to i_d, i_q and ω . Fig. 7 illustrates the phase plane. The negative values of the Lyapunov exponents viz. $-8.2999, -6.9554, -5.2624$, confirm the stability.

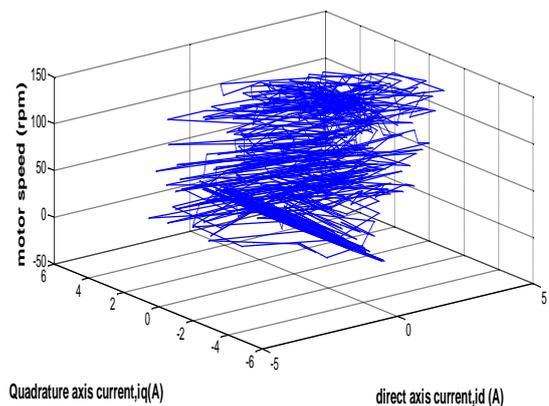


Figure 7 Phase plane trajectory

7. Conclusions

A Robust EKF estimator along with Fuzzy logic speed controller with an LMS estimate of the weighting coefficients has been successfully implemented in SIMULINK/MATLAB environment. The performance shows that the the change in speed is achieved within 0.5 S, with total stability in the control. The results illustrate the efficacy of the Bayesian approach for the elimination of outlier problem in the

EKF algorithm. The gradient descent algorithm used in the optimization of the combining of fuzzy rules included in the controller helps to catch up the sudden changes in speed very quickly within less than 500ms time. The Kalman filter proves its efficacy as an estimator without the knowledge of inputs, as a blind system identification approach. The controller performance has been tested for various step changes in speed, in the Simulink environment and established fast convergence and the stability at various speeds. .

8. Acknowledgement

The first author would like to thank the Institute of Human resource Development, Kerala, India for permitting her to do research work.

References

- [1] Wang Song, Shi Shuang-Shuang, Chen Chao, Yang Gang and Qu Zhi-jian, "Identification of PMSM based on EKF and Elman Neural network", *Proceedings of the IEEE International conference on Automation and Logistics* Shenyang, China August 2009
- [2] Bindu V, A Unnikrishnan, R Gopikakumari, "Adaptive fuzzy logic vector control of Permanent magnet synchronous motor with a robust Extended Kalman filter", *International journal on Industrial Electronics and control IJIEC (paper accepted)*
- [3] Bindu V, A Unnikrishnan, R Gopikakumari, "Fuzzy logic Vector control of PMSM with speed and rotor position estimation using a blind technique based on EKF", *International journal on Applied Engineering Research IJAER*, vol6, number7, 2011, pages 967-977
- [4] A Ting, E Theodorou, S Schaal, "Learning an Outlier Robust Kalman filter", *Lecture notes in Computer science*, 2003, vol4701, machine learning ECML 2007, pages 748-756
- [5] M N Uddin, T S Radwan, M A Rahman, "Fuzzy logic based position control of Permanent magnet synchronous motor", *Journal of Control and intelligent systems*, vol35, No.4, Nov 2007
- [6] Ying-Shieh Kung Ming-Hung Tsai, "FPGA-Based Speed Control IC for PMSM Drive With Adaptive Fuzzy Control", *IEEE Trans on power Electronics*, Nov 2007, vol22, issue 6 pages 2476-2486.
- [7] P Pillay, R Krishnan, "Modeling, Simulation and Analysis of Permanent magnet motor drives, Part 1: The Permanent magnet synchronous motor Drive",
- [8] D. Simon, Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches, *John Wiley & Sons*, 2006
- [9] J Wang, H Liu, "Novel intelligent sensorless control of Permanent magnet synchronous motor Drive", *The ninth International conference on Electronic measurements and Instruments, ICEMI2009*
IEEE Transactions on industry applications, vol25, No.2, March/April 1999
- [10] M. Barut, S. Bogosyan, M. Gokasa, "Speed sensor less direct torque control of IMs with rotor resistance estimation", *Science direct-Energy conversion and management*, July 2004 (electronic version)
- [11] F Parasilitary, R Petrella and M Tursini, "Sensorless speed control of a PM Synchronous motor based on sliding mode observer and Extended Kalman filter", *Industry applications conference, 2001, Thirty sixth IAS Annual meeting. Conference record of the 2001 IEEE*, vol.1, pp.533-540. Sep 4-October 2001
- [12] R. Krishnan, Electric Motor Drives, *Pearson Education, Singapore*, 2003.
- [13] Gilbert Foo, Saad Sayeef and M F Rahman, "An Extended Kalman filter for sensor less Direct torque controlled IPM synchronous motor drive", *2008 Australasian Universities power Engineering Conference (AUPEC'08)*
- [14] Dariusz Janiszewski, "Extended Kalman filter based speed sensor less PMSM control with load reconstruction", *IEEE Industrial Electronics, IECON 2006, 32nd Annual conference*
- [15] G Faten, S Lassaad, "Speed sensor less IFOC of PMSM bases on Adaptive Luenberger observer", *International journal of Electrical and Electronics Engineering 2:1:2009* technology and application, IEEE Press, 1997
- [16] W J Cao Binggang, Su Guang, "Fuzzy adaptive controller Design of Permanent magnet synchronous motor for electric vehicle", *2009 Third International conference on Intelligent information technology application*
- [17] Xin Xiaonan, He Li, "Chaotic model of Permanent magnet Linear Synchronous motor based on MATLAB", *2010 IEEE International conference ICEEE, Nov 2010, pages 1-4*

A Novel Session Based Dual Steganographic Technique Using DWT and Spread Spectrum

Tanmay Bhattacharya^{*}, Nilanjan Dey^{} and S. R. Bhadra Chaudhuri^{***}**

^{*}Asst. Professor Dept. of IT, JIS College of Engineering, Kalyani, West Bengal, India.

^{**}Asst. Professor Dept. of IT, JIS College of Engineering, Kalyani, West Bengal, India.

^{***}Professor, Dept of E&TCE, Bengal Engineering and Science University Shibpur, Howrah, West Bengal, India.

ABSTRACT

This paper proposed a DWT based Steganographic technique. Cover image is decomposed into four sub bands using DWT. Two secret images are embedded within the HL and HH sub bands respectively. During embedding secret images are dispersed within each band using a pseudo random sequence and a Session key. Secret images are extracted using the session key and the size of the images. In this approach the stego image generated is of acceptable level of imperceptibility and distortion compared to the cover image and the overall security is high.

Keywords - DWT, Session Based Key, Pseudo Random Sequence

I. INTRODUCTION

Steganography [1, 2, 3] is the process of hiding of a secret message within an ordinary message and extracting it at its destination. Anyone else viewing the message will fail to know that it contains secret/encrypted data. The word comes from the Greek word “*steganos*” meaning “covered” and “*graphei*” meaning “writing”.

LSB [4] insertion is a very simple and common approach to embedding information in an image in special domain. The limitation of this approach is vulnerable to every slight image manipulation. Converting image from one format to another format and back could destroy information secret in LSBs. Stego-images can be easily detected by statistical analysis like histogram analysis. This technique involves replacing N least significant bit of each pixel of a container

image with the data of a secret message. Stego-image gets destroyed as N increases. In frequency domain data can be made secret by using Discrete Cosine Transformation (DCT) [5, 8]. Main limitation of this approach is blocking artefact. Grouping the pixel into 8×8 blocks and transforming the pixel blocks into 64 DCT co-efficient each. A modification of a single DCT co-efficient will affect all 64 image pixels in that block. One of the modern techniques of Steganography is Discrete Wavelet Transformation (DWT) approach [6, 7]. In this approach the imperceptibility and distortion of the Stego image is acceptable and it is resistant to several attacks.

II. DISCRETE WAVELET TRANSFORMATION

The wavelet transform describes a multi-resolution decomposition process in terms of expansion of an image onto a set of wavelet basis function. The wavelet transform describes a multi-resolution decomposition process in terms of expansion of an image onto a set of wavelet basis functions. Discrete Wavelet Transformation has its own excellent space frequency localization properly. Applying DWT in 2D images corresponds to 2D filter image processing in each dimension. The input image is divided into 4 non-overlapping multi-resolution sub-bands by the filters, namely (LL1), (LH1), (HL1) and (HH1). The sub-band (LL1) is processed further to obtain the next coarser scale of wavelet coefficients, until some final scale “ N ” is reached. When “ N ” is reached, we’ll have $3N+1$ sub-bands consisting of the multi-resolution sub-bands (LLN) and

(LHX), (HLX) and (HHX) where “X” ranges from 1 until “N”. Generally most of the Image energy is stored in these sub-bands.

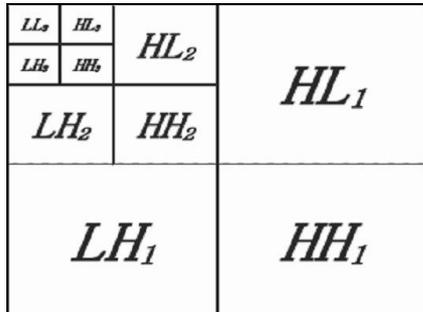


Fig.1 Three phase decomposition using DWT.

The Forward Discrete Wavelet Transform is very suitable to identify the areas in the cover image where a secret image can be embedded effectively due to its excellent space-frequency localization properties. In particular, this property allows the exploitation of the masking effect of the human visual system such that if a DWT co-efficient is modified, it modifies only the region corresponding to that coefficient. The embedding secret image in the lower frequency sub-bands (LL_x) may degrade the image significantly, as generally most of the Image energy is stored in these sub-bands. Embedding in the low-frequency sub-bands, however, could increase robustness significantly. In contrast, the edges and textures of the image and the human eye are not generally sensitive to changes in the high frequency sub-bands (HH_x). This allows the stego-image to be embedded without being perceived by the human eye. The compromise adopted by many DWT based algorithms, to achieve acceptable performance of imperceptibility and robustness, is to embed the secret image in the middle frequency sub-bands (LH_x) or (HL_x) and (HH_x). The Haar wavelet is also the simplest possible wavelet. Haar wavelet is not continuous, and therefore not differentiable. This property can, however, be an advantage for the analysis of signals with sudden transitions.

III. CODE DIVISION MULTIPLE ACCESS (CDMA) SPREAD-SPECTRUM TECHNIQUE

Spread-spectrum technique can be described as a method in which a signal generated in a particular bandwidth when deliberately spread in the frequency domain, results in a signal with a wider bandwidth. If distortion is introduced in this signal by some process such as noise or filtering which damages only certain bands of frequencies, the message will be still in a recoverable state. In spread spectrum communications, the signal energy inserted into any one frequency is too undersized to create a visible artefact and the secret image is scattered over a wide range of frequencies, that it becomes robust against many common signal distortions. Because of its good correlation properties, noise like characteristics, easier to generate and resistance to interference, Pseudo noise sequences are used for Steganography.

IV. PROPOSED ALGORITHM

Secret Image Hiding:

1. Cover image is decomposed into four sub bands (LL, LH, HL and HH) using DWT.
2. Two secret images are taken and converted into two different 1D Vectors.
3. Two different pseudo random 2D sequences are generated by the session based key.
4. Each HL and HH sub band of the cover image are modified separately using corresponding PN sequence depending upon the content of the corresponding secret 1D image vector to be embedded.
5. Four sub bands including two modified sub bands are combined to generate the stego image using IDWT.

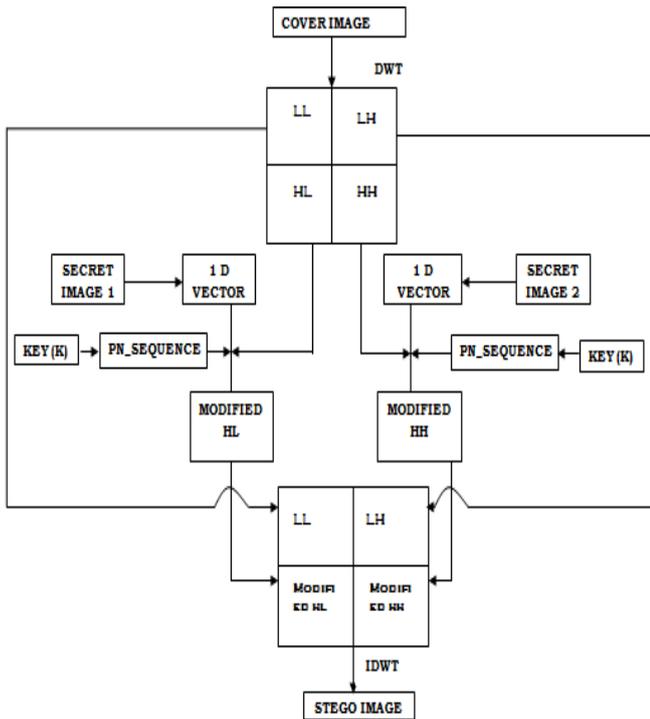


Fig 2: Image Hiding Process

Secret Image Extraction

1. Session key and Sizes of the secret images are sent to the intended receiver via a secret communication channel.
2. Secret images can recovered from the stego image using Correlation function and knowing the size of the secret image.
3. Extracted Secret Images are filtered to remove the unwanted signal.

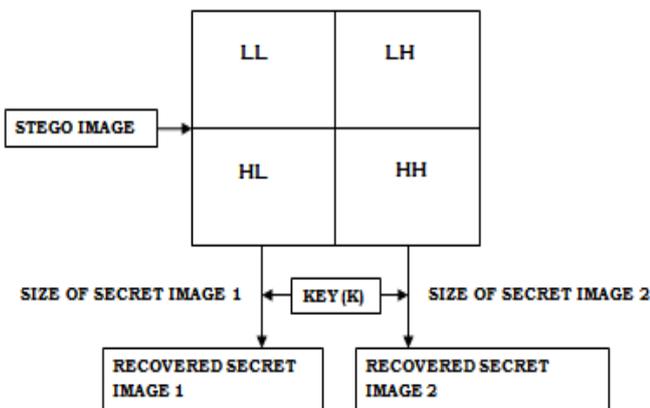


Fig 3: Image Extraction Process

V. EXPLANATION OF THE ALGORITHM

Secret Image Hiding

Using DWT the Cover image is decomposed into four sub bands (LL, LH, HL and HH).

Two binary images, Secret Image 1 and Secret Image 2 are taken and converted into 2 one dimensional vectors.

Two pseudo random sequences are generated using a session based key and the size of any sub bands of the cover image.

Each of the bits of the binary secret image1 and binary secret image2 are embedded in HL and HH sub-bands respectively depending upon the elements of the one dimensional vector and the pseudo random sequences. The general equation used to embed the secret image is:

$$I_s(x, y) = I(x, y) + k \times S(x, y) \dots\dots\dots (1)$$

In which $I(x, y)$ representing the selected DWT sub band of the cover image, $I_s(x, y)$ is the modified cover image, K denotes the amplification factor that is usually used to adjust the invisibility of the secret images in corresponding sub bands. $S(x, y)$ is the pseudo random sequences.

Taking all the sub bands including the modified HH and LH sub bands, stego image is obtained applying IDWT (Inverse Discrete Wavelet Transformation)

Secret Image Extraction procedure

The session key and the size of the secret images are provided to the intended receiver through a secret communication channel.

The sub-bands are selected into which the secret image was embedded after applying DWT on stego image. The pseudo random sequence (PN) is generated using the same session based key which was used in the secret image embedding procedure. The correlation between the selected stego sub-band and the generated pseudo random sequence is calculated. Each correlation value with the mean correlation value is compared. If the calculated value is greater than twice the mean, then the extracted watermark bit will be

taken as a 0, otherwise it is taken as a 1. The recovery process then iterates through the entire PN sequence until all the bits of the Secret image1 and secret image 2 have been recovered.

Filter is used on recovered secret images to remove unwanted signals.

VI. RESULTS



Fig. 4: Cover Image



Fig. 5: Secret Image-1



Fig. 6: Secret Image2



Fig. 7: Stego Image



Fig. 8: Recovered Secret Image -1



Fig. 9: Recovered Secret Image -2

Peak Signal to Noise Ratio (PSNR)

It measures the quality of a stego image. This is basically a performance metric and use to determine perceptual transparency of the stego image with respect to host image:

$$PSNR = \frac{MN \max_{x,y} P_{x,y}^2}{\sum_{x,y} (P_{x,y} - \bar{P}_{x,y})^2} \tag{1}$$

Where, M and N are number of rows and columns in the input image,

$P_{x,y}$ Is the original image and
 $\bar{P}_{x,y}$ Is the Stego Image.

PSNR between Cover Image and Stego Image is 27.3850 shown in Table1.

| | |
|-----------------|---------|
| Cover Image | PSNR |
| vs. Stego Image | 27.3850 |

Table 1

Correlation coefficient

After secret image embedding process, the similarity of original cover image x and stego images x' was measured by the standard correlation coefficient as follows:

$$Correlation = \frac{\sum (x - x')(y - y')}{\sqrt{(x - x')^2} \sqrt{(y - y')^2}} \tag{2}$$

Where y and y' are the discrete wavelet transforms of x and x'

Correlation between the secret image1 and recovered secret image1 after applying filter is 0.9381 and between the secret image2 and recovered secret image2 is 0.8870 shown in Table 2.

| Correlation between original secret image and recovered secret image | Image1 | Image 2 |
|--|--------|---------|
| | 0.9381 | 0.8870 |

Table 2

VII. CONCLUSION

In the proposed method second image is embedded in the HL sub band of the cover image. So there is a small visual change in between cover image and stego image. But due strong security aspects this small amount of imperceptibility is acceptable. This approach can be applied for colour image and audio Steganography also because DWT is applicable for any digital signal.

REFERENCES

- [1] N. F. Johnson and S. Katzenbeisser, "A survey of Steganographic techniques", in S. Katzenbeisser and F. Peticolas (Eds.): *Information Hiding*, pp.43-78. Artech House, Norwood, MA, 2000.
- [2] Lou, D. C. and Liu, J. L. 2002. "Steganography Method for Secure Communications". *Elsevier Science on Computers & Security*, 21, 5: 449-460.
- [3] J. Fidrich and M. Goljan, "Practical steganalysis of digital images-state of the art.", *Proc. SPIE Photonics West*, Vol. 4675, pp. 1-13, San Jose, California, Jan. 2002.
- [4] Chan, C. K. and Cheng, L. M. 2003. Hiding data in image by simple LSB substitution. *Pattern Recognition*, 37:469-474.
- [5] Iwata, M., Miyake, K., and Shiozaki, A. 2004. "Digital Steganography Utilizing Features of JPEG Images", *IEICE Transfusion Fundamentals*, E87-A, 4:929-936.
- [6] Po-Yueh Chen* and Hung-Ju Lin, "A DWT Based Approach for Image Steganography", *International Journal of Applied Science and Engineering* 2006. 4, 3: 275-290
- [7] Ali Al-Ataby and Fawzi Al-Naima, "A Modified High Capacity Image Steganography Technique Based on Wavelet Transform", *The International Arab Journal of Information Technology*, Vol. 7, No. 4, October 2010
- [8] Blossom Kaur, Amandeep Kaur, Jasdeep Singh, "Steganographic Approach for Hiding Image in DCT Domain", *International Journal of Advances in Engineering & Technology*, July 2011.

Communication Enhancement Using Socket Programming

Monika Sharma¹, Prashant Soni²

* (Faculty, AIIT, Amity University, India)

*(Student, Amity University, India)

Abstract

As a java programmer one might face networking using socket programming. A network-based system consists of a server, client, and a media for communication and for such communication we use "Sockets".

Sockets are interfaces that can "plug into" each other over a network. Once so "plugged in", the programs so connected communicate. A socket is one end-point of a two-way communication link between two programs running on the network. Socket classes are used to represent the connection between a client program and a server program. Such programming is called socket programming.

The java.net package of "JAVA" provides two classes—Socket and ServerSocket, that implement the client side of the connection and the server side of the connection, respectively.

Keywords –Client, Java.net, Sockets, Socket Programming, Socket (Class), Server and ServerSocket (Class).

1. Introduction

There are many codes developed for socket programming which were previously being implemented but some drawbacks like slow speed of execution, they were not user friendly, uneven flow of data, etc hindered their best possible use and I have tried to patch them.

As far as security is concerned, the 'class' file of java plays its role and thus provide much security

for code on being copied to any other operating system.

1.1 Java TCP Programming

The programming model of TCP communication in Java, rely completely on the sockets and ports. Because TCP is a stream protocol, it allows to send arbitrary amount of data rather rely on class to encapsulate data within TCP packets.

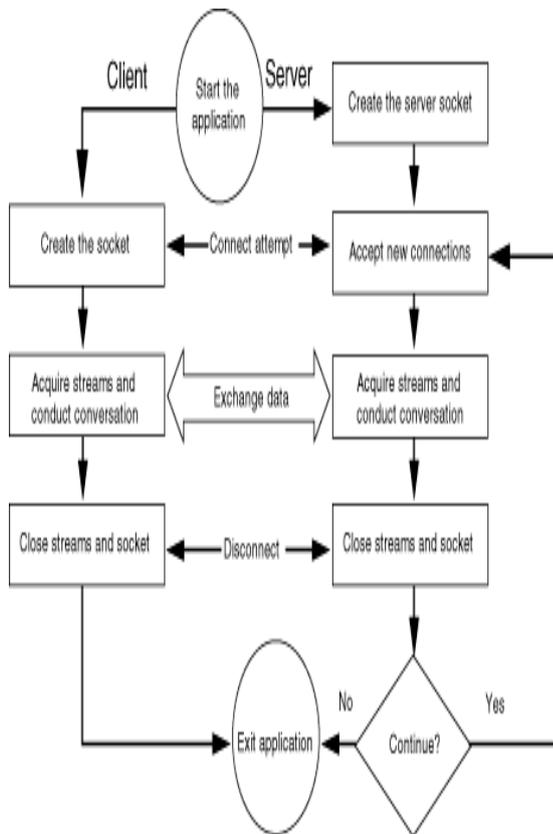
1.2 Sockets

Java programs communicate through a programming abstraction called socket. A socket is one end-point of a two-way communication link between two computers (or programs) running on a network. A socket is bound to a port number so that the TCP layer can identify the application that data is destined to be sent. Mostly, applications do not care how data is actually transmitted in the network. Applications identify the address of the peer entity and then use sockets interface to read and write data from and to the peer.

Sockets include the implementation of network and transport layer protocols providing applications with a simple read/write interface. Because sockets are just programming abstractions for network protocols, the other side of the connection does not have to use them. Sockets don't use any additional communication mechanism other than that provided by the encapsulated protocol^[1].

1.3 Ports

The mechanism, commonly used by network protocols, and particularly by the Internet transport layer protocols, is port addressing^[2]. For example, tcp:23 port number is assigned for FTP, tcp:80 for HTTP, etc.



1.5.1 This figure illustrates how can client and server interact with each other

Typically, integer numbers are used to identify different ports. In order to contact a network service, it is therefore necessary to provide both the IP address of its host, as well as port number it is using.

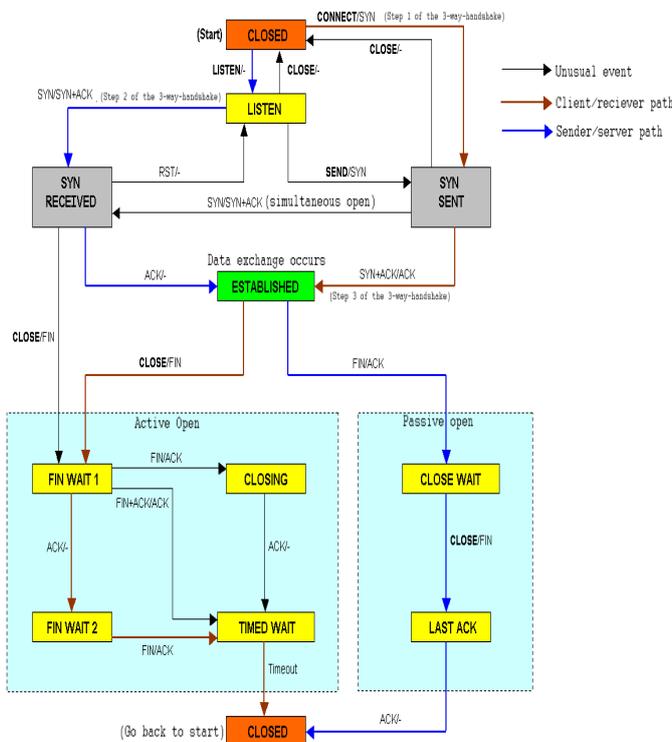
1.4 Connection Establishment

TCP uses a 3 way handshake. A Server makes a passive open by call bind(). The Client initiates an active open by calling connect(). For the connection to be established, the client sends a SYN packet to the server. The Server replies with SYN/ACK packet and finally the client replies with ACK packet [3].

1.5 Connection Termination

TCP uses a 4 way handshake to close the connection. When an endpoint wants to close, it sends out a FIN packet, the other side then replies with an ACK packet [3].

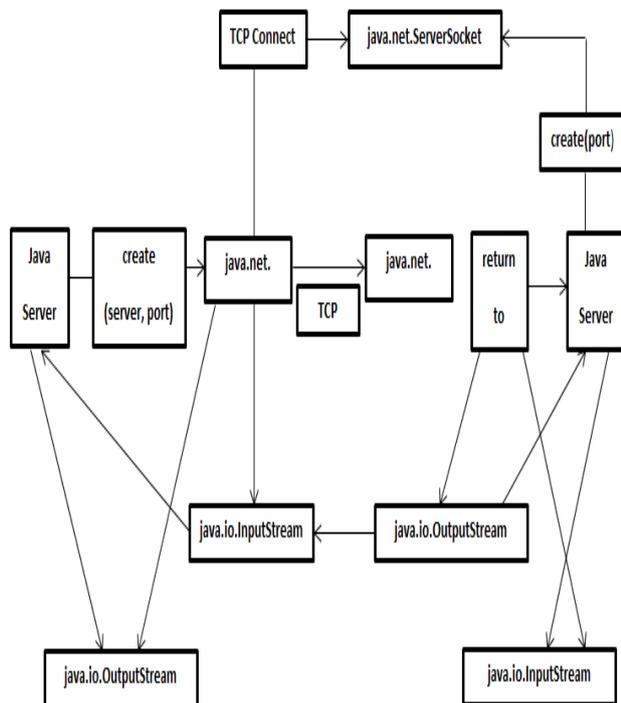
A connection is “half-open” when one side has close the connection. The other side is still free to send. A situation can occur where one side closes the connection and then reopens it immediately, so any lost packets that now arrive will not belong to this “new” connection and thus TCP we need to insure these packets do not get mixed in with the new packets. So the “TIME WAIT” state allows a connection to remain open long enough for such packets to be removed from the network. This state usually lasts for about 2 times the “round-trip”, some implementation hardcode the default value to be anywhere from 30 to 120 seconds. We can use the netstat utility to see TCP/IP states.



1.5.2 This figure illustrates how a three way and a four way handshake takes place.

1.6 Package java.net

The Java language supports TCP programming through the java.net.Socket and java.net.ServerSocket classes. Java clients connect to TCP servers by creating instances of the java.net.Socket class. Similarly, Java servers listen for java clients by creating java.net.ServerSocket class [3]. Connections are configured through the methods of these two classes. Actual network communications, however, are performed using the java.io [4] package streaming classes.



1.6.1 The figure illustrates the main operations of the Java streaming socket classes.

The Java TCP server, on the right, accepts TCP connections by creating a **java.net.ServerSocket** instance bound to a particular port.

When a **ServerSocket** establishes a connection with a TCP client, it creates a **java.net.Socket** instance encapsulating that connection and returns it to the server program. The **java.net.Socket** object returned is bound to an ephemeral ports number that is different from the one the **ServerSocket** is listening to. The server retrieves the socket's input and output streams and effects communication by implementing some protocol.

On the client side, TCP connections are established through instances of **java.net.Socket** associated with a TCP server on the given host and port. Once the client **Socket** has established a TCP connection, retrieves the socket's input and output stream, effects communication.

2 Interpretation and Implementation

2.1 Sending Data Through A Socket

Once an application program has established a socket, it can use the socket to transmit data.

While sending the data we check that whether the data is null or not. If the data is not null then we may proceed to send data to stream through socket at receiver's end ^[5].

2.2 Receiving Data through a Socket

For receiving the data also we shall use socket and in same way as data was sent to stream, we will receive it. The data received through socket is read and printed on output screen.

2.3 Client-Side TCP Programming

The **java.net.Socket** of a **java.net** package implements client side of a two-way connection between Java program and another program on the network. By using this class instead of relying on native code, Java program can communicate over network in platform independent fashion ^[6].

2.3.1 Creating a Socket

In order to establish a connection with a network server one must have the address of the server's host, and port number to which the server is bound. The **java.net.Socket** class provides a constructor, which takes an IP address and port number and attempts to establish a connection. The signature of this constructor is as follows –

`Socket(InetAddress, int port)` throws `IOException`;

The constructor returns only after the connection has been established, that is, once the TCP three-way handshake has been completed.

2.4 Server-side TCP Programming

In order to accept network connections a Java program must create an instance of **java.net.ServerSocket**. Server sockets are not directly used to perform any network communication. Instead, they act as factories that create a **java.net.Socket** object for every incoming TCP communication request. Programs create the server socket, bind to a specific port on one or more interfaces, and then invoke the blocking **accept()** method ^[7].

2.4.1 Creating ServerSocket

The basic **ServerSocket** constructor takes a single argument, the TCP port number used in binding. If the constructor returns without throwing an exception then the server socket has successfully bound

to the requested TCP port. The constructor may fail due to an I/O error, or due to a security error. The signature of the constructor is –

ServerSocket(int port) throws IOException, SecurityException;

2.4.2 Accepting Sockets

The main task of server socket is to receive incoming connection requests and generate a **java.net.Socket** object that encapsulates each request. Incoming connections are queued until the program retrieves them one at a time by invoking the accept() method. The accept() method takes no arguments, and returns the next connection in the queue.

2.5 Terminating Server Socket

A server socket may be terminated simply by invoking the no – argument close() method.

Closing the server socket will not affect connections that have already been returned by accept () invocation. If the accept () method is invoked on a closed socket then a **java.net.Socket** exception will be thrown with a message indicating that the socket has been closed. The signature is as follows –

void close() throws IO Exception;

3 Conclusion

Developing network applications is made possible in Java by using sockets, threads, RMI, clustering, and Web services. These technologies allow for the creation of portable, efficient, and maintainable large and complex Internet applications. The java.net package provides a powerful and flexible set of classes for implementing network applications.

Typically, programs running on client machines make requests to programs on a server machine. These involve networking services provided by the transport layer. The most widely used transport protocols on the Internet are TCP (Transmission control Protocol) and UDP (User Datagram Protocol). TCP is a connection-oriented protocol providing a reliable flow of data between two computers. It is used by applications such as the World Wide Web, e-mail, IP, and secure shell.

Sockets provide an interface for programming networks at the transport layer. Using sockets, network communication is very much similar to per-

forming file I/O. A socket is an endpoint of a two-way communication link between two programs running on the network. The source and destination IP address, and the port numbers constitute a network socket.

Two key classes from java.net package used in creation of server and client programs are **ServerSocket**, which represents a server socket, and **Socket**, an instantiation of which performs the actual communication with the client.

4 Future Work

In future, for further optimization, I am trying to implement security constraints on my code like ‘Cryptography’ for enhanced two way communication.

References

Articles:

- [1] <http://www.buyya.com/java/Chapter13.pdf>
- [2] <http://www.scribd.com/doc/29858538/Java-TCP-Programming>
- [3] http://www.pcvr.nl/tcpip/tcp_conn.htm
- [4] <http://download.oracle.com/javase/1.5.0/docs/api/> (Java API(Application Programming Interface) 5.0)
- [5] <http://edn.embarcadero.com/article/31995>

Books:

- [6] Herb Schildt, *Java 2 The Complete Reference* (4th edition, Osborne/MacGraw-Hill, 2001)
- [7] Kathy Sierra & Bert Bates, *Head First Java* (2nd edition, O’Reilly Media, 2005)

Enhanced channel access mechanism for VoIP services in IEEE 802.11 Networks

B.Suvarna

Vignana university

B.Lalunaick

narsarapotea Engineering college

Abstract:

Transmitting voice through IP data network can provide significant cost savings. However if not managed properly, voice quality can degrade due to data network congestion. VoIP is voice over an Internet Protocol (IP) based network.

In this paper we will investigate QoS indications (for voice packets), Improvement of NS2 network simulator in order to simulate statistical QoS in 802.11 MAC. The IEEE 802.11e Standard has been introduced recently for providing Quality of Service (QoS) capabilities in the emerging wireless local area networks. This 802.11e introduces a contention window based that is Enhanced Distribution Channel Access (EDCA) technique that provides a prioritized traffic to guarantee minimum bandwidth needed for time critical applications. However this EDCA technique resets statistically the contention window of the mobile station after each successful transmission. This static behavior does not adapt to the network state hence reduces the network usage and results in bad performance and poor link utilization whenever the demand for link utilization increases. For that purpose a new adaptive differentiation technique has been proposed for IEEE 802.11e wireless local area networks that take into account the network state before resetting the contention window.

To improve the QoS of Voice over Internet Protocol services we proposed a new traffic for VoIP. The performance of the proposed technique and proposed traffic is evaluated and compared with the original IEEE802.11a technique. Preliminary results show that the proposed adaptive technique enhances the channel utilization and increases throughput

Keywords: Voice over Internet Protocol (VoIP), QoS, WLANs, DCF, enhanced DCF.

1.Introduction

VoIP stands for Voice over Internet Protocol. As the term says VoIP tries to transfer voice (mainly human) using IP packets over the Internet. Voices over IP (VoIP) applications are gaining an ever increasing popularity in the Internet community, favored by the massive deployment of wireless

access technologies. For instance, more than eighty million users have already subscribed to Skype, the most popular VoIP commercial application for personal use, roughly 10% of which are estimated to be simultaneously online at any time. While it is not clear whether VoIP will ultimately replace traditional telephony, its massive diffusion may act as the main driving factor for the actual deployment of Quality of Service (QoS), both in the Internet backbone and in the (wired or wireless) access segments. For this reason, using VoIP as a test case in the performance evaluation of new QoS components, such as (to name a few) scheduling, resource reservation, admission control, traffic policing, traffic engineering, etc., has become a common practice. Unlike classic data applications, in which easily quantifiable, data-related performance metrics (e.g., throughput and mean packet delay) most often represent meaningful evaluations, the actual performance of VoIP applications depends on user perception (a concept often referred to as Quality of Experience, QoE). For this reason, the ITU-T has established a computational model, called the Emodel, which defines a quality factor - the so-called R score — to capture the effect of mouth-to-ear delay and losses in packet-switched networks. The R score can then be mapped to the Mean Opinion Score (MOS), which in

turn can be converted into subjective quality levels (e.g. “good”, “poor”). Despite this, assessing the VoIP performance through measures taken at the IP level – rather than taking into account the user perception – is often the norm in QoS literature. However, it can be shown that a sound assessment of VoIP quality has to take into account several factors which extend beyond the IP level. For instance, playout buffers, which come as part of a VoIP application, play a crucial role: packets that are successfully delivered within a given deadline at the IP level can in fact be delayed or dropped at the playout buffer.

How does VoIP work? Before sending the voice across the network, VoIP digitalizes it in data packets, sends them and reconverts them to voice at destination. Why do we convert it to the digital format? Digital format can be better controlled: we can compress it, route it and convert it to a new better format, and so on. In addition, digital signals are more noise tolerant than analog ones.

Overview on a VoIP connection:

- First, we use ADC to convert analog voice to digital signals (bits) - This is made by hardware, typically by card integrated ADC.
- Now the bits have to be compressed in a good format for transmission: there are a number of protocols, for example PCM, Pulse Code Modulation, Standard ITU-T G.711. The most important demand from such protocols is to convert digital data to a standard format that could be quickly transmitted.
- Here we have to insert our voice packets in data packets using a real-time protocol. VoIP data packets are packed in RTP (Real-Time Transport Protocol) packets, which are inside UDP-IP packets. VoIP doesn't use TCP because it is too heavy for real time applications, so instead a UDP datagram is used.
- However, UDP has no control over the order in which packets arrive at the destination or how long it takes them to get there. Both of these are very important to overall voice quality and conversation quality. RTP solves the problem enabling the receiver to put the packets back into the correct order and not wait too long for packets that have either lost their way or are taking too long to arrive (we don't need every single voice packet,

but we need a continuous flow of many of them and ordered).

- We need a signaling protocol to call users: ITU-T H323 does that. This protocol allows a variety of elements talking each other: terminals, clients that initialize VoIP connection, Multipoint Control Units (MCUs) to provide conference and more... This protocol allows not only VoIP but also video and data communications.
- At the receiver we have to disassemble packets, extract data, then convert it to analog voice signals and send it to sound card (or phone).
- All that must be done in a real time fashion because we cannot wait for too long for a vocal answer.

2. Background

VoIP is voice over an Internet Protocol (IP) based network. All networks will be supporting IP. There are two ways of looking at VoIP: regulatory/business and technical. We are going to address the technology. The regulatory and business perspective will provide a framework by which VoIP will be provided. However, the regulatory and business view is far too complex to discuss in a white paper. As a service, voice is a basic necessity. Despite the preponderance of email, people prefer to talk to one another rather than email one another. Declining minutes of use in the wireline network is due to the existence of wireless communications and email. As a mass market service, voice is the basic service of all services. Without voice a telecommunications service provider is not meeting the needs of all of its customers.

The Internet was not originally designed to carry audio communications. In fact the Internet protocol could not meet the exacting requirements of the voice service customer. Once an ISP is capable of providing voice it will be able to take advantage of its position as an information services provider to the user and provide all services (including voice) to the user. At one time, VoIP was provided as a best effort service just as other Internet services had been. The Internet Protocol is a “best effort” protocol.

In general there are business and technical benefits to deploying an IP network. The business benefits are:

- ❖ Reduced long distance costs
- ❖ Lower network costs
- ❖ More enhanced services – Voice over IP is just one of the services.

The technical benefits are:

- ❖ Less bandwidth for more calls
- ❖ More efficient use of network resources
- ❖ Distributed network intelligence

The network signaling protocol of the Internet is TCP/IP. The Transmission Control Protocol/Internet Protocol (TCP/IP) protocol suite was originally used for and still is used for the internetworking of Local Area Networks (LANs). All of the signaling protocols used in the Internet are part of the TCP/IP protocol suite. The Transport layer in the TCP/IP suite is comprised of two protocols; the TCP and the UDP. TCP stands for Transmission Control Protocol. UDP stands for User Datagram Protocol. The TCP (Transmission Control Protocol) performs the transport layer functions of the Internet Protocol. The UDP (User Datagram Protocol) is a connectionless function that is normally used by database lookup applications.

Although originally designed for data services, the Internet can also support real-time traffic such as voice and video. The technology of voice over Internet Protocol (VoIP), also known as Internet telephony, IP telephony, or packet voice, enables real-time voice conversations over the Internet. It has attracted much interest from academia and industry because of the following facts:

- VoIP has much lower cost than traditional telephone service.
- The universal presence of IP makes it convenient to launch VoIP applications.
- There is increasing demand for networks to interact with end users having real-time data, voice, and video images, leading to the requirement for integrated voice, data, and video services.
- The emerging digital signal processing (DSP) and voice coding/decoding techniques make VoIP more and more mature and feasible. Therefore, VoIP is anticipated to offer a viable alternative to traditional public switched telephone network (PSTN).

To provide person-to-person (instead of place-to-place) connections anywhere and anytime, the Internet is expected to penetrate the wireless domain. One very promising wireless network is the wireless local area network (WLAN), which has shown the potential to provide high-rate data services at low cost over local area coverage. Working in the license-exempt 2.4 GHz industrial, scientific, and medical (ISM) frequency band, the IEEE 802.11b WLAN offers a data rate up to 11 Mb/s, while IEEE 802.11a WLAN and European Telecommunications Standard Institute (ETSI) HIPERLAN/2 can support data rates

up to 54 Mb/s at the 5 GHz frequency band. As a wireless extension to the wired Ethernet, WLANs typically cover a small geographic area, in hotspot local areas where the traffic intensity is usually much higher than in other areas. The promising VoIP technology and wide deployment of WLANs are expected to drive the application of voice over WLAN (VoWLAN), which will experience a dramatic increase in the near future. Figure 1 shows a typical VoWLAN system where voice conversation happens through the access point (AP). At the sender, the analog voice signal is compressed and encoded by a codec. After inclusion of the Real-Time Transport Protocol (RTP)/User Datagram Protocol (UDP)/IP headers during the packetization procedure at the transport and network layers, voice packets are transmitted over the networks and finally to the receiver end. At the receiver, a playout buffer is usually used to alleviate the effect of delay jitter. Then the receiver applies depacketization and decoding to recover the original voice signal. One major challenge for VoWLAN is quality of service (QoS) provisioning. Originally designed for high-rate data traffic, WLANs may experience bandwidth inefficiency when supporting delay-sensitive and low-rate voice traffic. Hence, it is essential to enhance the QoS support capability of current WLAN standards, such as the most popular IEEE 802.11 standard.

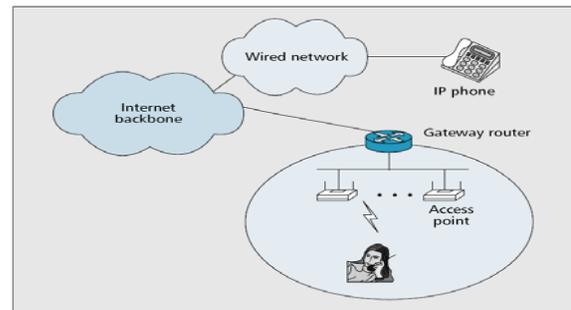


Figure 1. The architecture for VoIP over WLAN

3. Overview of 802.11 - 802.11a, 802.11b and 802.11e

802.11 -- refers to a family of specifications developed by the IEEE for wireless LAN technology. 802.11 specify an over-the-air interface between a wireless client and a base station or between two wireless clients. The IEEE accepted the specification in 1997. 802.11 defines physical and MAC layers and

provides 1 or 2 Mbps transmission in the 2.4 GHz band using either frequency hopping spread spectrum (FHSS) or direct sequence spread spectrum (DSSS).

802.11b (also referred to as 802.11 High Rate or Wi-Fi) -- an extension to 802.11 that applies to wireless LANs. 802.11b was a 1999's ratification to the original 802.11 standard, allowing wireless functionality comparable to Ethernet. Most WLANs deployed today use 802.11b technology, which operates in the 2.4 GHz band and supports a maximum theoretical data rate of 11 Mbps, with average throughput falling in the 4 Mbps to 6 Mbps range. In a typical office environment, its maximum range is 75 meters (250 feet) at the lowest speed, but at higher speed its range is about 30 meters (100 feet). Minimizing interference can be difficult because 802.11b uses only three non-overlapping channels. 802.11b uses only DSSS. So, its advantages are: it allows multiple connections to a remote network, data transfer and mobility. But, as we can see here, 802.11b isn't perfect at all. Actually, 802.11b has three major problems: limited bandwidth, interference from other devices and also it doesn't allow any Quality of Service (QoS).

802.11a -- an extension to 802.11 that applies to wireless LANs. Operating in the 5 GHz band, 802.11a supports a maximum theoretical data rate of 54 Mbps, but more realistically it will achieve throughput somewhere between 20 Mbps to 25 Mbps in normal traffic conditions. In a typical office environment, its maximum range is 50 meters (150 feet) at the lowest speed, but at higher speed, the range is less than 25 meters (75 feet). 802.11a has four, eight, or more channels, depending on the country. 802.11a uses an orthogonal frequency division multiplexing encoding (OFDM) scheme rather than FHSS or DSSS. In general, if we want high performance and minimal radio frequency interference, then 802.11a is the way to go (but without QoS!)

4.Channel accessing mechanisms in 802.11 a and 802.11b.

The IEEE 802.11 WLAN (both a and b) have two different channel accessing mechanisms, namely, the distributed coordination function (DCF) and point coordination function (PCF). DCF is based on the

carrier sense multiple access with collision avoidance (CSMA/CA) channel accessing mechanism, while PCF is based on the polling technique. The DCF operation mode consists of two techniques for packet transmission. The default scheme is a two-way handshaking technique where a positive acknowledgement is transmitted by the destination station upon successful reception of a packet from a sender station. Another scheme involves a four-way handshaking technique known as request to send/clear to send mechanism (RTS/CTS). By this scheme, the sender first sends RTS to reserve the channel before its transmission, and upon receiving CTS from the receiver, the normal packet transmission and the ACK response proceeds. On the other hand, for the PCF operation mode, stations are polled in turn, and the station with a packet pending for transmission sends the packet upon being polled. In IEEE 802.11 networks, the DCF mode is the fundamental channel access method and coexistence between DCF and PCF is required. The period in which the system operates in PCF mode is called contention free period (CFP), while the period in which the system operates in DCF mode is called contention period (CP). Moreover, using just PCF presents the following inefficiency: If every wireless station connected to an AP are polled regardless of whether it has data to transmit or not may result in considerable polling overhead. This overhead may be reduced by maintaining a dynamic polling list at the AP. A station with data to transmit asks the AP to enroll to this list and after some idle time the AP deletes it from the list. In this sense, DCF is still needed in addition to PCF, in order to provide the stations a way to send the enrollment requests to the AP.

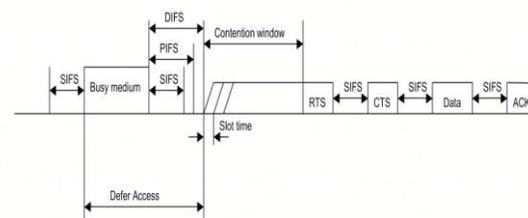


Figure 2. MAC Layer

802.11e -- Is an enhanced version of 802.11, currently under IEEE development. It keeps most of technical parameters of its predecessor, but has very significant quality: it provides Quality of Service (QoS) support for LAN applications, which will be critical for

delay-sensitive applications such as Voice over Wireless IP (VoWIP). The standard will provide classes of service with managed levels of QoS for data, voice, and video applications. It introduces the concept of hybrid coordination function (HCF) for the MAC mechanism. HCF is upward compatible DCF and PCF, in the same time providing QoS stations with prioritized and parameterized QoS access to the wireless medium.

EDCF and EPCF -- HCF provides two different means of supporting QoS. First there is the extension of the widely deployed distributed coordination function (DCF) that makes use of CSMA/CA. DCF provides coordination, but it doesn't support any type of priority access of the wireless medium. The enhanced DCF (EDCF) mechanism adds four levels of statistical access priority, enabling the separation of frames into different priority levels. Each level corresponds to an individual prioritized output queue. Each output queue contends for a transmission opportunity (TXOP). The minimal specified idle duration before starting a frame transmission (ICF – inter frame spaces) is different for each specific queue: SIFS (short IFS) is used by acknowledgement packets. PIFS (point coordination function IFS) is used by the AP to take control of the channel and start CFP. DIFS is used by data packets and so on. The backoff computation is also different for the individual queues. Contention window is increased after each collision. EDCF provides statistical priority only. It does not guarantee that low priority frames will always wait until all higher priority frames are transmitted. The second 802.11e QoS mechanism is an extension of PCF of the original 802.11 standard. This method uses a QoS-aware point coordinator, called hybrid coordinator (HC). The AP usually plays this role. The HC uses its higher channel access priority to allocate transmission rights (TXOPs) to wireless stations to transmit QoS data so that the predefined delivery priority, service rate and delay are satisfied. The wireless station may sent a TXOP request to the HC either while in EDCF mode, or during another TXOP granted to it or in a special CCI interval (controlled contention interval) when contention occurs only among QoS stations wishing to get a TXOP. During a TXOP the station may initiate multiple frame exchange sequences. This gives EPCF the flexibility to support bursty QoS traffic. EPCF inherently provides hard QoS guaranties.

802.11e allows Quality of Service (QoS), while the original 802.11a and 802.11b do not .

The proposed approach is based on adapting the values of CW depending on the channel congestion level. In IEEE 802.11e the value of CW is incremented whenever a station fails to transmit due to a collision. This would imply that when the channel is highly congested CW would acquire values distant from CW_{min} and close to CW_{max}. Similarly, when the channel is free, CW values would be close to CW_{min} and distant from CW_{max}. Hence, it is feasible to estimate the channel congestion level by taking into consideration the current value of CW. We use a very simple approach to estimate this level. In this approach, we start from the fact that CW value ranges in the interval [CW_{min}, CW_{max}], then we compute its relative distance $(CW_{current} - CW_{min})$ compared to the maximum distance $(CW_{max} - CW_{min})$ as an indication for channel congestion level. It follows that the estimated link congestion ratio in the proposed Adaptive scheme can be written as:

$$Ratio = \frac{(CW_{current} - CW_{min})}{(CW_{max} - CW_{min})}$$

In the proposed scheme, the ratio is weighted as follows.

$$ratio = weight \times \frac{(CW_{current} - CW_{min})}{(CW_{max} - CW_{min})}$$

For instance, the weight of the ratio would be very small if current channel estimate is used in a transmission that occurred several minutes ago. However the ratio would be highly weighted if the difference in time between estimation and transmission is of the order of milliseconds. To obtain some preliminary simulation results, the weight was fixed in this paper to a value of 0.9. Indeed, the weight converged to this value after several tests. This is due to fact that video streaming is characterized by transmission occurring at very small time intervals

The CW value of the proposed adaptive scheme, CW_{new} can be given then as follows

$$CW_{new} = weight \times \frac{(CW_{current} - CW_{min})^2}{(CW_{max} - CW_{min})} + CW_{min}$$

The ratio is a normalized value ranging from 0 to 1 that reflects the weighted degree of channel contention. This ratio would take a value close to 0 whenever the channel is free. Therefore $CW_{current}$ would have a value close to CW_{min} and distant from CW_{max} . The value of this ratio would be close to 1 whenever the channel is congested. Therefore $CW_{current}$ would have a value distant from CW_{min} and close to CW_{max} . Multiplying this ratio by the factor $(CW_{current} - CW_{min})$ and adding the result to CW_{min} would result in a value bounded by $[CW_{min}, CW_{max}]$. This value of CW_{new} would be a good representation of the backoff timer value needed for transmission for the current traffic priority taking into account the current network conditions.

5. Simulation scenario

The simulation topology of this scenario is simple. It consists of 8 mobile nodes: 4 source nodes and 4 destination nodes. Each node is transmitting with a different priority. Node 1 is given a higher priority than Node 2, which is given also a higher priority than Node 3. Node 3, in its turn, is given a higher priority than Node 4. Each source is a Constant Bit Rate source over UDP (User Datagram Protocol). The size of a transmitted packet is 512 bytes. Transmission rate of a node is 600Kbps. We assumed that the nodes are in transmission range at a constant distance of 195 m. The simulation time lasted for 80 sec.

To model voice traffic the simulations use ITU-T G.729 standard. G.729 is supported widely in VoIP products. In G.729, the voice is encoded at the rate of 8 kbps and with 20 or 40 bytes payload size in a packet. The voice quality can be degraded compared to another widely used standard, G.711, because the compression in G.729 can be lossy. However G.729 requires less bandwidth. The payload size is 20 bytes. With packet overhead, the rate required is 26.4 kbits/s

6. Simulation results

In this section we present simulation results, which are meant as a proof of concept of how the contributed simulation framework can be exploited for a sound and simple performance evaluation of VoIP applications in ns-2. We therefore purposefully set up a very simple networking environment.

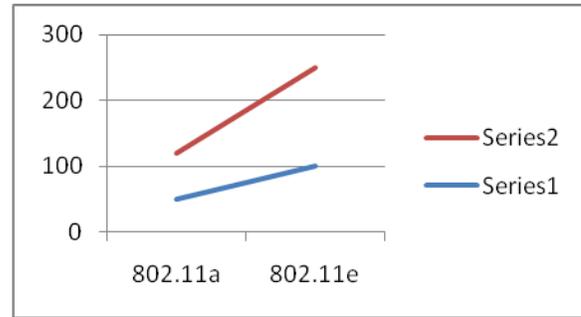


Figure 4. packet transfer rate

Bit rate

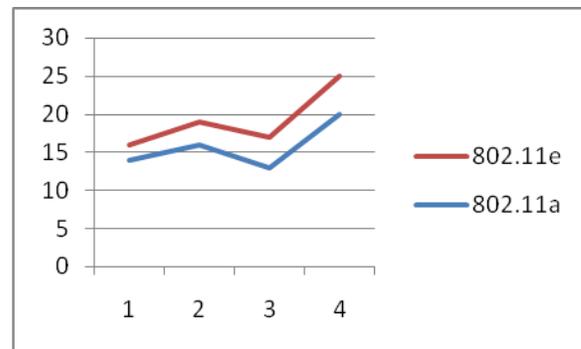


Figure 4. packet loss rate

7. Conclusion

In this project the performance of the IEEE 802.11a and IEEE 802.11e systems have been evaluated. We have also proposed a new adaptive differentiation technique for resetting the value of the contention window after each successful transmission. The proposed adaptive technique takes into account the current level of link utilization when resetting such value. We have performed several simulations, for different Scenarios, using NS-2, to evaluate the proposed technique compared to IEEE 802.11a and IEEE 802.11e. During the evaluation, we have focused on three parameters: bit rate, end-to-end packet delay. And the results reveal that the proposed traffic gives better results than IEEE802.11a.

8. References

- [1]. P. Chatzimisios, V. Vitas as, A. Boucouvalas and M. Tsoulfa, Achieving performance enhancement in IEEE 802.11 Weans by using DIDD back off mechanism, International Journal of Communication Systems, vol. 20, no. 1, pp. 23-41, 2007.
- [2]. A. Bacioccola, C.Cicconetti, G.SteaUser level performance evaluation of VoIP using NS-2, Computer Networks Group, 2007
- [3] I. -H. Lin and J.-Y. Pan, Throughput and delay analysis of a novel back off algorithm for IEEE 802.11 Weans, Proc. Wireless Telecommunications Symposium, pp. 85- 90, 2005.
- [4]. L. Autoscore, and M. L.Lobina. Play out buffering in ip telephony: a survey discussing problems and approaches .IEEE Commun surveys Tut, vol. 8, 2006
- [5] Mamun I. Abu-Tair, Geyong Min, Qiang Ni, Hong Liu "Adaptive Medium Access control for VoIP services using IEEE 802.11 Networks", IEEE 2008
- [6]Mamun abu Tair.Geyong Min,Qiang Ni,Hong Liu "Adaptive medium access control for VoIP services using IEEE 802.11 Networks

Parametric Approach to Analysis of Aluminum alloy Helical gear for High Speed Marine Applications

B.Venkatesh¹, V.Kamala², A.M.K. Prasad³

*(Dept of Mechanical Engineering, Vardhaman College of Engineering, Hyderabad, A.P, India.)

** (Dept of Mechanical Engineering Methodist College of Engineering & Technology, Hyderabad, A.P, India.)

*** (Dept of Mechanical Engineering University College of Engineering, Osmania University, Hyderabad, A.P, India.)

ABSTRACT

In the present era of sophisticated technology, gear design has evolved to a high degree of perfection. The design and manufacture of precision cut gears, made from materials of high strength, have made it possible to produce gears which are capable of transmitting extremely large loads at extremely high circumferential speeds with very little noise, vibration and other undesirable aspects of gear drives. The gears are used for power transmission and also in marine and aerospace applications; where accuracy and power transmission capacity are the requirements. High speed gears (of the order of 30,000rpm) have to be of high quality to maintain contact and keep contact stresses within the limits. The analytical investigation is based on modified Lewis stress formula. The present work deals with the optimization of design and analysis of helical gear with aluminum alloy for high-speed marine applications.

Keywords - Beam strength, bending stress, dynamic strength of tooth, Gear design, high speed helical gear.

I. INTRODUCTION

The motion from one shaft to another shaft may be transmitted with belts, ropes and chains. These methods are mostly used when the two shafts are having long center distance. But if the distance between the two shafts is very small, then gears are used to transmit motion from one shaft to another. In case of belts and ropes, the drive is not positive. There is slip and creep that reduces velocity ratio. But gear drive is a positive and smooth drive, which transmit velocity ratio. Gears are used in many fields and under a wide range of conditions such as in smaller watches and instruments to the heaviest and most powerful machineries like lifting cranes. Gears are most commonly used for power transmission in all the modern devices. They have been used extensively in the high-speed marine engines. There is a great deal of researches on gear analysis. Generally their major concerns are on the analysis of gear stresses, transmission errors, dynamic loads, noise, and failure of gear tooth, which are very useful for optimal design of gear set. They have used various approaches and

means to attain their main objectives. The first systematic studies in gear dynamics started in the 1920s by A.A.Ross and E.Buckingham . The basic concern in their studies was the prediction of tooth dynamic loads for designing gears to operate at high speeds. Helical gears are the modified form of spur gears, in which all the teeth are cut at a constant angle, known as helix angle, to the axis of the gear, where as in spur gear, teeth are cut parallel to the axis. The following are the requirements that must be met in the design of gear drive: the gear teeth should have sufficient strength, so that they will not fail under static and dynamic loading during normal running conditions. The gear teeth should have clear characteristics so that their life is satisfactory, the use of space and material should be economical. The alignment of the gears and deflections of the Shafts must be considered, because they affect the Performance of the gears. The lubrications of the gears must be satisfactory.

II. NOMENCLATURE

$[\sigma_b]$ = Design Bending stress in N/mm²

E= Young's modulus in N/mm²

[Mt] = design torque in N-mm

σ_b = Bending stress N/mm²

β = Helix angle in degrees

F_d = Dynamic tooth load in N

F_b = Beam Strength of the gear tooth

FD = Design tooth load

m_n = Normal Module

Y_v = Lewis Form factor

b =Face width in mm

DESIGN METHODOLOGY

Design of helical gear is based on "AGMA" Procedure: Beam Strength of helical gear tooth, according to Lewis equation is given by

$$F_b = \left[\frac{F_t}{b} \right] \cdot \pi m_n \cdot y_v$$

Virtual number of teeth $Z_v = [Z/\cos^3 \beta]$

Design tooth load F_D,

$$F_D = F_t \times K_s \times C_v = \frac{F_t \times K_s \times C_v}{v}$$

Bending stress $\sigma_b = 0.7*(i+1) [M_T] / (a b m_n Y_V)$

Buckingham equation for dynamic load acting on gear

$$F_d = F_t + \frac{21v \sqrt{Cb \cos^2 \beta + F_t} \cos \beta}{21v + \sqrt{Cb \cos^2 \beta + F_t}}$$

$$\text{Wear Strength of tooth load } F_w = \frac{d_1 \cdot b \cdot Q \cdot K_w}{\cos^2 \beta}$$

III. RESULT AND DISCUSSIONS

It is aimed at arriving optimum values of bending stress, dynamic tooth load, beam strength to achieve low cost of manufacturing for Aluminum alloy by carrying out analysis under different operating parameters.

The effect of gear ratio, face width, helix angle, normal module on bending stress for Aluminum alloy

The variation of bending stress for different input variables are shown in figures.1(a) – (d). The fig 1(a) shows the relation between bending stress and gear ratio. The helix angle, face width, speed and normal module except gear ratio are kept constant. When the gear ratio is increased from 4 to 8, the corresponding bending stress remained constant. The fig 1(b) shows the relation between bending stress and face width. The helix angles, gear ratio, Speed, normal module except face width are kept constant. When face width is increased from 41 to 49, the corresponding bending stress is observed to decrease linearly from 562 kgf/cm² to 465 kgf/cm². The fig 1(c) shows the relation between bending stress and helix angle. The face width, gear ratio, speed and normal module except helix angle are kept constant. When helix angle is increased from 15° to 35°, the corresponding bending stress was observed to decrease from 560 kgf/cm² to 470 kgf/cm². The fig 1(d) shows the relation between bending stress and normal module. The face width, gear ratio, speed and helix angle except normal module are kept constant. When normal module is increased from 16mm to 24mm, the corresponding bending stress was observed to decrease from 560 kgf/cm² to 240 kgf/cm².

The effect of gear ratio, face width, helix angle, normal module on Dynamic tooth load for Aluminum alloy

The variation of Dynamic tooth load for different input variables are shown in figures. 2(a)-(d). The fig 2(a) shows the relationship between Dynamic tooth load and gear ratio. The helix angle, face width, speed and normal module except gear ratio are kept constant. When gear ratio is increased from 4 to 8, the corresponding Dynamic tooth load remained constant. The fig 2(b) shows the relationship between Dynamic tooth load and Face width. The Helix angles, gear ratio, Speed, normal module except face width are kept constant. When face width is increased from 41 to 49, the corresponding Dynamic tooth load is constant. The fig 2(c) shows the relationship between Dynamic tooth load and Helix angle. The face width, gear ratio, speed and normal module except Helix angle are kept constant. When Helix angle is increased from 15° to 35°, the corresponding Dynamic tooth load decreased from 3340kgf to 2830kgf.

The fig 2(d) shows the relationship between Dynamic tooth load and Normal module. The face width, gear ratio, speed and Helix angle except Normal module are kept constant. When Normal module is increased from 16mm to 24mm, the corresponding Dynamic tooth load decreased from 3350kgf to 2250kgf.

The effect of gear ratio, face width, helix angle, normal module on Beam Strength for Aluminum alloy

The variation of Beam Strength for different input variables are shown in figures. 3(a) – (d). The fig 3(a) shows the relationship between Beam Strength and gear ratio. The helix angle, face width, speed and normal module except gear ratio are kept constant. When gear ratio is increased from 4 to 8, the corresponding Beam Strength remained constant. The fig 3(b) shows the relationship between Beam Strength and Face width. The Helix angles, gear ratio, Speed, normal module except face width are kept constant. When face width is increased from 41 to 49, the corresponding Beam Strength increased from 3025kgf to 3625kgf. The fig 3(c) shows the relationship between Beam Strength and Helix angle. The face width, gear ratio, speed and normal module except Helix angle are kept constant. When Helix angle is increased from 15° to 35°, the corresponding Beam Strength remained constant. The fig 3(d) shows the relationship between Beam Strength and Normal module. The face width, gear ratio, speed and Helix angle except Normal module are kept constant. When Normal module is increased from 16mm to 24mm, the corresponding Beam Strength increased from 3030kgf to 4580kgf.

IV. DISCUSSIONS

In Parametric study of helical gear made of Aluminum alloy material, the variation of parameters viz bending stress, dynamic tooth load, beam strength for different modules (i.e 16, 18, 20, 22 & 24) and different face width (i.e 41,43,45,47 & 49) The gear ratio (i) = 4,5,6,7 & 8 and helix angle (β) = 15, 20, 25, 30 & 35 are respectively kept constant. If the module is increased, the corresponding bending stress, dynamic tooth load was observed to decrease. However, beam strength were observed to gradually increase.

Optimum parameters for maximum bending stress: The effect of gear ratio, face width, helix angle, and normal module on optimum bending stress for aluminum alloy is carried out. If the helix angle, face width, speed and normal module except gear ratio are kept constant and the gear ratio is increased, the corresponding bending stress remained constant. Next the helix angle, gear ratio, speed, normal module except face width are kept constant and the face width is increased, the corresponding bending stress decreases linearly. The face width 41cm, corresponding to maximum bending stress is taken as constant value. The face width, gear ratio, speed and normal module except helix angle are kept constant and helix angle is increased, the corresponding bending stress was observed to decrease. The helix angle 15°, corresponding to maximum bending

stress is taken as constant. The face width, gear ratio, speed and helix angle except normal module are kept constant and normal module increases, the corresponding bending stress was observed to decrease. The normal module 16mm, corresponding to maximum bending stress is taken as constant.

Optimum parameters for maximum dynamic tooth load:

The effect of gear ratio, face width, helix angle, and normal module on optimum dynamic tooth load is carried out. If the helix angle, face width, speed and normal module except gear ratio are kept constant and gear ratio is increased, the corresponding dynamic tooth load remained constant. When the helix angles, gear ratio, speed, normal module except face width are kept constant and face width is increased, the corresponding dynamic tooth load remained constant. When the face width, gear ratio, speed and normal module except helix angle are kept constant and helix angle is increased, the corresponding dynamic tooth load decreases. The helix angle 15° , corresponding to maximum dynamic tooth load remains constant. The face width, gear ratio, speed and helix angle except normal module are kept constant and normal module is increased, the corresponding dynamic tooth load decreases. The normal module 16mm, corresponding to maximum dynamic tooth load is taken as constant.

Optimum parameters for maximum beam strength:

The effect of gear ratio, face width, helix angle, normal module on optimum beam strength is carried out. If the helix angle, face width, speed and normal module except gear ratio are kept constant and gear ratio is increased, the corresponding beam strength remained constant. The helix angles, gear ratio, speed, normal module except face width are kept constant and face width is increased, the corresponding beam strength found to increase. The face width 49cm, corresponding to maximum beam strength is taken as constant. The face width, gear ratio, speed and normal module except helix angle are kept constant and helix angle is increased, the corresponding beam strength remained constant. The face width, gear ratio, speed and helix angle except normal module are kept constant and normal module is increased, the corresponding beam strength found to increase. The normal module 24mm, corresponding to maximum beam strength is taken as constant.

Manganese has been known to be an alloying element of Al alloys that contributes to uniform deformation. Recently, it was found that as the manganese content increases over 0.5% in such aluminum alloys, both yield and ultimate tensile strength increase significantly without decreasing ductility. The added manganese forms a manganese dispersoid, this dispersoid has an incoherent structural relationship with respect to the matrix, FCC, in retarding the motion of dislocations that increase strength. Once the dislocation is blocked by the dispersoid, it tends to change the slip system by means of cross-slip. This cross-slip allows the deformation to maintain uniformly good ductility. Adding manganese to aluminum alloys not only enhances tensile strength but also significantly improves low-cycle fatigue resistance. Corrosion resistance is also measurably improved by the

addition of manganese. After extrusion, the recrystallization is also retarded so that a very small grain size is maintained, contributing to an improvement in the mechanical properties.

The addition of magnesium to aluminum increases strength through solid solution strengthening and improves their strain hardening ability. These alloys are the highest strength non-heat-treatable aluminum alloys and are, therefore, used extensively for structural applications. The alloys are produced mainly as sheet and plate and only occasionally as extrusions. The reason for this is that these alloys strain harden quickly and, are, therefore difficult and expensive to extrude.

V. CONCLUSIONS

Present day competitive business in global market has brought increasing awareness to optimize gear design. Current trends in engineering globalization require results to comply with various normalized standards to determine their common fundamentals and those approaches needed to identify "best practices" in industries. This can lead to various benefits including reduction in redundancies, cost containment related to adjustments between manufacturers for missing part interchangeability, and performance due to incompatibility of different standards. From the study of effect of various parameters (viz. bending stress, dynamic tooth load, beam strength) on the optimum design of helical gears for marine applications, the induced bending stresses are much lower than those of the results obtained theoretically. Also the bending stresses are much lower than the design stresses, thus the design is safe from the structural point of view. It is observed that the induced bending stresses are less than that of the theoretical calculations. Aluminum alloy reduces the weight up to 55-67% compared to other materials like steel. Weight reduction is a very important criterion, in order to minimize the unbalanced forces setup in the marine gear system, there by improves the system performance. The helical gear parameters that constitute the design are found to be safe from strength and rigidity point of view. Hence Aluminum alloy may be best possible material for marine gear in the high speed applications.

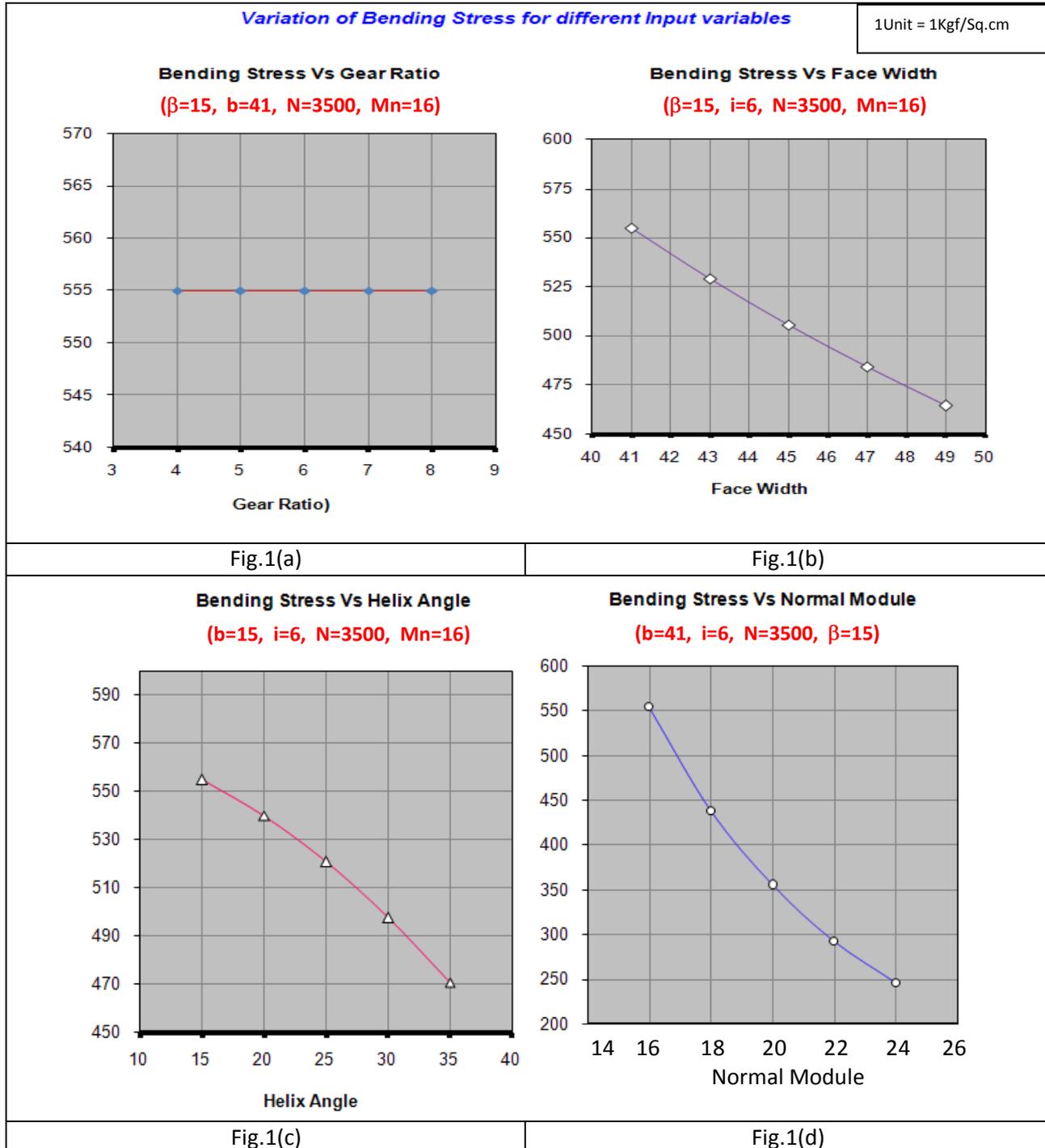
VI. ACKNOWLEDGMENTS

First author is grateful to Dr.N.Sambasivarao, Principal Vardhaman College of Engineering, Hyderabad and Dr.G.V.Rao, Head Mechanical Engineering Department for their valuable support. He is also thankful to the management of the institute for the encouragement, support, and co-operation during the entire work.

REFERENCES

- [1] J.A. Wright, et al, Design, development and application of new, high-performance gear steels, *Gear technology* (2010),pp 46 – 53
- [2] Rao, C.M., and Muthuveerappan G., Finite Element Modeling and Stress Analysis of Helical Gear, Teeth, *Computers & structures*, 49,pp.1095-1106, 1993.

- [3] Marappan, S. and Venkataramana, 2004, *ANSYS Reference Guide.*, CAD CENTRE, India.
- [4] Cheng, Y., and Tsay C.B., Stress analysis of Helical Gear set with Localized Bearing Contact, *Finite Element in Analysis and Design*, 38, pp.707-723, 2002.
- [5] Song He, Rajendra Gunda, Rajendra Singh., 2007 Inclusion of Sliding Friction in Contact Dynamics Model for Helical Gears, *ASME Journal of Mechanical Design*, Vol. 129, pp.48-57.
- [6] PSG, 2008. *Design data*, Kalaikathir Achchagam publishers, Coimbatore, India.

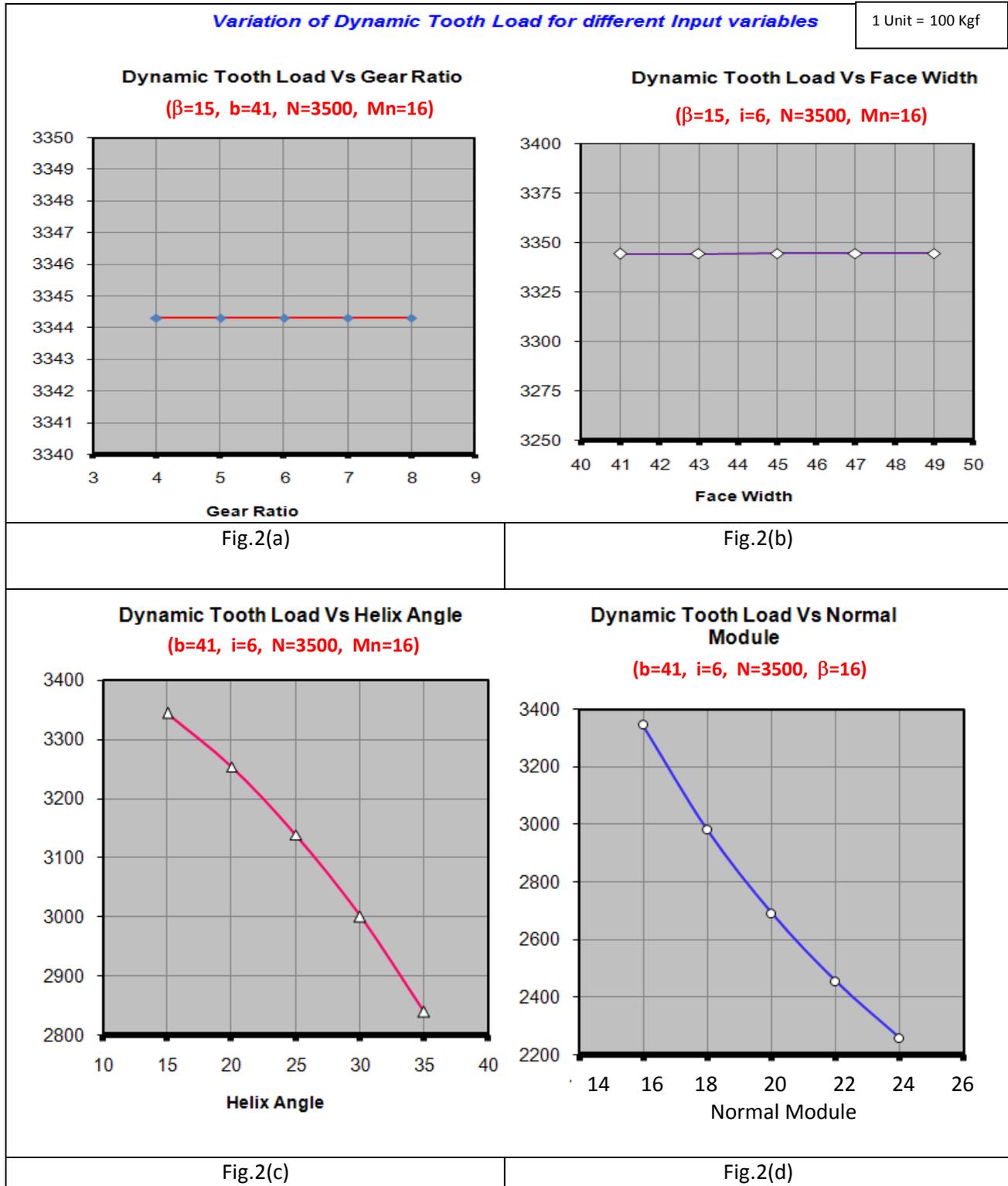


[7] Preloaded Gearing for high speed application Lazar Chalik, DE-Vol 88, *Power transmission & Gearing conference ASME* 1996

[9] R.E. Sanders, Technology Innovation in aluminum Products, *The Journal of The Minerals*, 53(2):21–25, 2001

[8] Darle W. Dudley, *Handbook of practical gear design.*, 1954

[10] [http://: www.matweb.com](http://www.matweb.com)



Variation of Beam Strength for different Input variables

1 Unit = 100 Kgf

Beam Strength Vs Gear Ratio
($\beta=15$, $b=41$, $N=3500$, $Mn=16$)

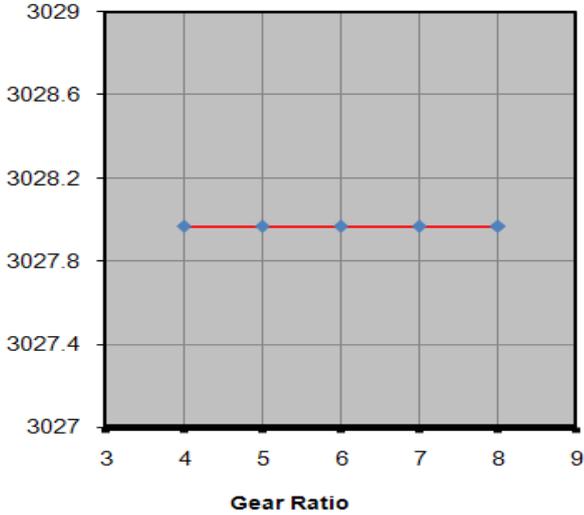


Fig.3(a)

Beam Strength Vs Face Width
($\beta=15$, $i=6$, $N=3500$, $Mn=16$)

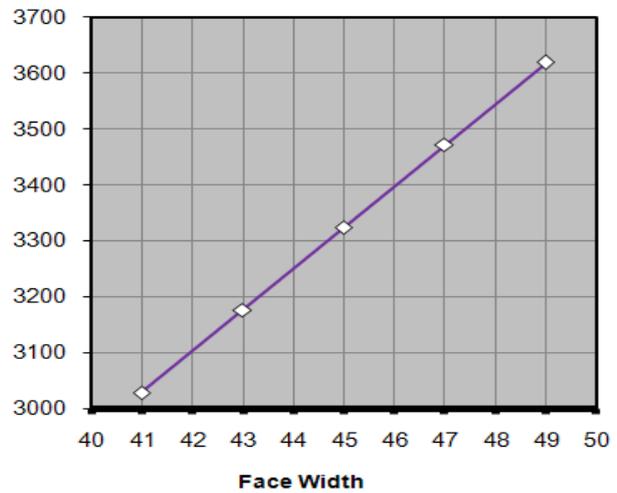


Fig.3(b)

Beam Strength Vs Helix Angle
($b=41$, $i=6$, $N=3500$, $Mn=16$)

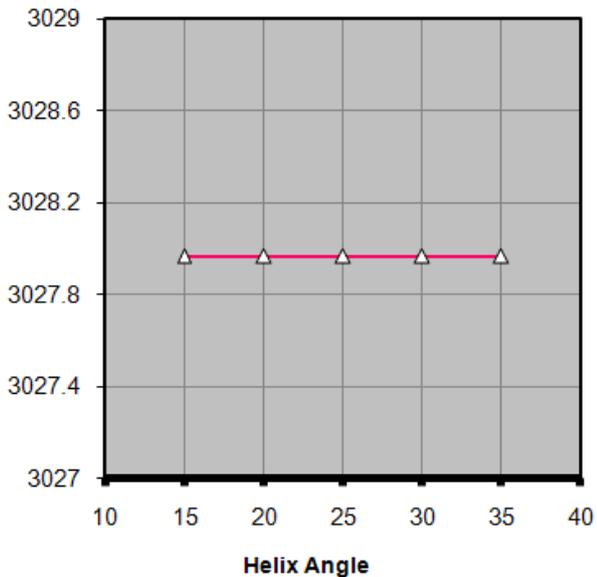


Fig.3(c)

Beam Strength Vs Normal Module
($b=41$, $i=6$, $N=3500$, $\beta=15$)

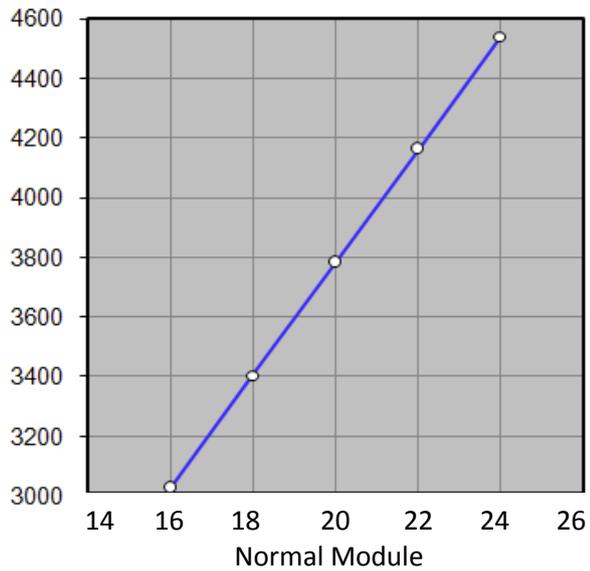


Fig.3(d)

Comparative Study of Partial Encryption of Images and Video

Ms. A.Anto Steffi^{*}, Mr. Dipesh Sharma^{**}

^{*}(Department of Computer Science & Engg., RIT Raipur, Chhattisgarh, India)

^{**} (Department of Computer Science & Engg., RIT Raipur, Chhattisgarh, India)

ABSTRACT

The traffic of digital images and video has grown rapidly in the internet. Security becomes important for several applications like military image database, confidential video conferencing, medical images, etc. Several techniques have been developed for textual data but are not appropriate for images and video with huge amount of file size. Partial encryption is a recent approach to reduce the encryption time of images and video in distributed network. Partial encryption scheme encrypts a portion of compressed bit stream. In this paper we compare and classified various proposed partial encryption schemes for images and video.

Keywords - Compression, JPEG, MPEG, Partial Encryption

1. INTRODUCTION

The increased popularity of multimedia applications has demanded a certain level of security. In some applications, it is relevant to hide the content of a message when it enters an insecure channel. The initial message prepared by the sender is then converted into cipher text prior to transmission. The process of converting plain text into cipher text is called encryption. The encryption process requires an encryption algorithm and a key. The process of recovering plain text from cipher text is called decryption. Because common encryption methods generally manipulate an entire data set, most encryption algorithms tend to make transfer of information more costly in terms of time and sometimes bandwidth. Traditionally, an appropriate compression algorithm is applied to the multimedia

data and its output is encrypted by an independent encryption algorithm. This process must be reversed by the receiver.

Unfortunately, the processing time for encryption and decryption is a major factor in real-time image communication. In addition, the processing time required for compression and decompression of an associated image data is important. Encryption and decryption algorithms are too slow to handle the tremendous amount of data transmitted. One difference between text data and image data is that the size of image data is much larger than the text data. The time is a very important factor for the image encryption. We find it at two levels, one is the time to encrypt, the other is the time to transfer images. To minimize the time, the first step is to choose a robust, rapid and easy method to implement cryptosystem. The other important criteria concerns the method of compression is that to decrease the size of images without loss of image quality [1]. One possible solution is a system of partial encryption, encrypting only the smallest portion of the data that makes the entire data set unusable. Partial encryption is a recent approach to reduce the computational requirements for huge volumes of multimedia data. Partial encryption is currently an important research area. We will start with a classification and brief description of the proposed schemes in order to identify some of the related problems.

2. PARTIAL ENCRYPTION

The encryption algorithms, which have been originally developed for text data, are not suitable for securing many real time algorithms, which have been originally developed for text data, are not suitable for

securing many real time multimedia applications because of large data sizes. Software implementations of ciphers are usually too slow to process image and video data in commercial systems. Hardware implementations, on the other hand, add more cost to service providers and consumer electronics device manufacturers. Recent trend is to minimize the computational requirements for secure multimedia distribution by “partial encryption” where only parts of the data are encrypted.

2.1 Partial Encryption Schemes for Images

- Cheng and Li, 2000

Cheng and Li [2] proposed partial encryption methods that are suitable for images compressed with two specific classes of compression algorithms:

- a) quad tree image compression algorithms

It allows the encryption and decryption time to be significantly reduced without affecting the compression performance of the underlying compression algorithm. In this scheme, the compression output is partitioned into two parts; one is important and other is unimportant parts. Important parts provide a significant amount of information about original data, whereas remaining part called unimportant parts may not provide much information without important parts. Encryption will only perform for important parts. A significant reduction in encryption and decryption time is achieved when the relative size of important part is small. This scheme is not tunable as static parameters are encrypted. High visual degradation can be achieved only with image having high information rate. As encryption is performed after compression, so no impact is observed on compression efficiency. Encryption ratio can vary from 14% to 50%. Brute force attack is possible for low information images where quad tree structure is very simple. So the security level of this scheme is low.

- b) Wavelet compression based on zero trees

In general, wavelet compression algorithms based on zero trees transmit the structure of the zero tree with the significant coefficients. The SPIHT algorithm, for example, transmits the significance of the coefficient sets that

correspond to trees of coefficients. Among the many different types of bits generated by the SPIHT algorithm, the proposed partial encryption scheme encrypts only the significance information related to pixels or sets in the two highest pyramid levels in addition to the parameter n that determines the initial threshold.

- Droogenbroeck and Benedett, 2002

- a) Proposed selective encryption methods for raster images and JPEG images. In their method the DC coefficients are not ciphered because they carry important visible information and they are highly predictable. Moreover, in their approach the compression and encryption stages are separated and that requires an additional operating cost [3].

- b) This method is proposed for uncompressed image, which applies to a binary image, consist in mixing image data and a message (key) that has the same size as the image: a XOR function is sufficient when the message is only used once. A generalization to gray level images is straightforward: Encrypt each bit plane separately and reconstruct gray level image. With this approach no distinction between bit planes is introduced although the subjective relevance of each bit plane is not equal. The highest bit planes exhibit some similarities with the gray level image, but the least significant bit planes look random. Because encrypted bits also look random, the encryption of least significant bit planes will add noise to the image. The advantage of least significant bits is that plaintext attacks are harder on random like data. It is preferable to encrypt bits that look most random. This scheme is tunable. Very high visual degradation can be achieved by encrypting 4 to 5 bit planes. This technique is used for uncompress image so no impact is observed on compression efficiency. In this scheme encryption ratio vary from 50 to 60%. It is fast as XOR operation takes less time. It is not robust against cryptanalysis attack. So, security level is low.

- Podesser, Schmitdt and Uhl, 2002

In selective bit plane encryption using AES is proposed. Several experiments were conducted on 8 bit grayscale images, and the main results retained are following: 1. encrypting only the

MSB is not secure; a replacement attack is possible 2. Encrypting the first two MSBs gives hard visual degradation, and 3. Encrypting three bit planes gives very hard visual degradation. This scheme is not tunable as fix number of bits are encrypted. For 8 bits per pixel uncompressed image, hard visual degradation (of 9 dB) can be observed for a minimum of 3MSB bits encrypted. This scheme is intended for uncompressed data. Encryption can increase data size so it is not compression friendly. In this scheme encryption is performed before compression, so it is format compliant. At least 3 bit planes over 8 (more than 37.5%) of the bit stream have to be encrypted using AES to achieve sufficient security even when a secure cipher is used (AES), the selective encryption algorithm proposed is vulnerable to replacement attacks. This attack does not break AES but replaces the encrypted data with an intelligible one. It is worth to note that visual distortion is a subjective criterion and does not allow to measure security as illustrated in this example. Security level of this technique can be scaled as medium [4].

- Pommer and Uhl, 2003

The authors proposed wavelet packet based compression instead of pyramidal compression schemes in order to provide confidentiality. Header information of a wavelet packet image coding scheme that is based on either a uniform scalar quantizer or zero trees is protected: it uses AES to encrypt only the sub band decomposition structure. In this approach the encoder uses different decomposition schemes with respect to the wavelet packet sub band structure for each image. It is based on AES encryption of the header information of wavelet packet encoding of an image, this header specifies the sub band tree structure. These decomposition trees are encrypted and have to be present at the decoder to be able to reconstruct the image data properly. The advantage in comparison to other selective encryption approaches is that the amount of necessary encryption is extremely small since only header information, and no visual data, needs to be processed. It is not tunable. The encrypted content cannot be viewed without decryption. The sub band tree is pseudo randomly generated. This adversely impacts the compression efficiency. It is not format compliant. The encrypted part represents a very

small fraction of the bit stream. It is not secure against chosen plaintext attack. Because statistical properties of wavelet coefficients are preserved by the encryption, then the approximation sub band can be reconstructed. This will give the attacker the size of the approximation sub band (lower resolution) and then neighbouring sub bands can be reconstructed since close sub bands contain highly correlated coefficients [5].

2.2 Partial Encryption Schemes for Video

- Meyer and Gadget, 1995

This methodology is proposed for MPEG videos. This method uses traditional encryption methods RSA or DES in CBC mode to encrypt MPEG video stream. It implements 4 level of security. (i) Encrypting all stream headers. (ii) Encrypting all stream headers and all DC and lower AC coefficients of intracoded blocks. (iii) Encrypting I-frames and all I-blocks in P- and B frames. (iv) Encrypting all the bit streams. The number of I blocks in P or B frames can be of the same order as the number of I blocks in I frames. This reduces considerably the efficiency of the selective encryption scheme. Encryption ratio may vary based on which parameters are encrypted. Encrypting only headers have very less encryption ratio. But encrypting all the bit streams have 100% encryption ratio. Speed of this methodology again varies based on traditional algorithm in use such as DES or RSA and number of parameters that are encrypted. Many security levels can be obtained. Encrypting only stream headers is not sufficient since this part is easily predictable. But encrypting all the bit streams can provide high security. Detailed cryptanalysis of this methodology is not defined. A special encoder and decoder are required to read unencrypted SEC MPEG stream. The encoder proposed is not MPEG compliant [6].

- Spanos and Maples, 1995

Aegis mechanism is proposed. It encrypts intraframes, video stream header and the ISO 32 bits end code of the MPEG stream using DES in CBC mode. Experimental results were conducted by the authors showing the importance of selective encryption in high bit rate video transmission to achieve acceptable end-to-end delay. It is also shown that full encryption

creates bottleneck in high bit rate distributed video applications. Agi and Gong showed that this algorithm has low security since encrypting of only I-frames offer limited security because of the intercorrelation of frames; some blocks are intracoded in P and B frames. Furthermore, P and B-frames are highly correlated when they correspond to the same I-frame. They also underlined that it is unwise to encrypt stream headers since they are predictable and can be broken by plaintext-cipher text pairs. Alatter and Al-ragib, apparently unaware of Agi and Gong work, stressed the same security leakage. Encryption is performed after compression, thus no impact is observed on the compression efficiency. The resulting bit stream is not MPEG compliant [7].

- Shi and Bhargava, 1998

The authors [8] proposed video encryption algorithm (VEA) which uses a secret key to randomly change the signs of all DCT coefficients in an MPEG stream. It is fast as it operates on a small portion of original video. It is more efficient than DES algorithm because it only selectively encrypts a small number of bits of the MPEG compressed video and selected bit is only XORed one time with the corresponding bit of the secret key. VEA does not protect from plaintext attack provided the attacker knows the original video image (plaintext and cipher text). The authors present a new version of VEA reducing computational complexity; it encrypts the sign bits of differential values of DC coefficients of I-frames and sign bits of differential values of motion vectors of Band P-frames. This type of improvement makes the video playback more random and more non viewable. When the sign bits of differential values of motion vectors are changed, the directions of motion vectors change as well. In addition, the magnitude of motion vectors change, making the whole video very chaotic. Modified VEA encrypt DC coefficients of I frame, and leave AC coefficients of I frames unchanged. Thus it significantly reduces encryption computations. Because DC coefficients of I frames are differentially encoded, changing a few sign bits of differential values of DC coefficients will affect many DC coefficients during MPEG decoding. MPEG's differential code of DC coefficients and motion vectors increase the difficulty to break MVEA

encrypted videos. The first version of VEA [21] is only secure if the secret key is used once. Otherwise, knowing one plaintext and the corresponding cipher text, the secret key can be computed by XORing the DCT sign bits. Both versions of VEA are vulnerable to chosen plaintext attacks; it is feasible to create a repetitive/periodic pattern and then compute its inverse DCT. The encryption of the image obtained will allow us to get the key length and even compute the secret key by chosen-plaintext attack.

- Shi, Wang and Bhargava, 1999

A new version of the modified VEA presented is proposed, called real time video encryption algorithm (RVEA) [9]. It encrypts selected sign bits of the DC coefficients and/or sign bits of motion vectors using DES or IDEA. It selects at most 64 sign bits from each macro block. RVEA achieves the goal of reducing and bounding its computation time by limiting the maximum number of bits selected. The differential encoding of DC coefficients and motion vectors in MPEG compression increases difficulty of breaking RVEA encrypted videos. If the initial guess of a DC coefficient wrong, it is very difficult to guess the following DC values correctly.

- Wu and Kuo, 2001

It [10] is based on a set of observations, the authors point out that energy concentration does not mean intelligibility concentration. Indeed, they discussed the technique proposed by Tang. They show that by fixing DC values at a fixed value and recovering AC coefficients (by known or chosen plaintext attacks), a semantically good reconstruction of the image is obtained. Even using a very small fraction of the AC coefficients does not fully destroy the image semantic content. The authors argued that both orthogonal transform-based compression algorithms followed by quantization and compression algorithms that end with an entropy coder stage are bad candidates to selective encryption. They investigate another approach that turns entropy coders into ciphers. They propose two schemes for the most popular entropy coders: multiple Huffman tables (MHTs) for the Huffman coder and multiple state index (MSI) for the QM arithmetic coder.

MHT: The authors propose a method using multiple Huffman coding tables. The input data stream is encoded using multiple Huffman tables. The content of these tables and the order that they are used are kept secret as the key for decryption. In the proposed system, instead of training thousands of Huffman coding tables, it only train and obtained four different Huffman tables. Then, thousands of different tables can be derived using a technique called Huffman tree mutation. Gillman and Rivest showed that decoding a Huffman coded bit stream without any knowledge about the Huffman coding tables would be very difficult. However, the basic MHT is vulnerable to known and chosen plaintext attacks.

MSI: The arithmetic QM coder is based on an initial state index; the idea is to select 4 published initial state indices and to use them in a random but secret order. Unlike Huffman coding with a fixed and pre defined Huffman tree, the QM coder dynamically adjusts the underlying statistical model to a sequence of received binary symbols. It is very difficult to decode the bit stream without the knowledge of the state index used to initialize the MQcoder. A little effect on compression efficiency is observed. This is due to multiple initializations of the QM coder due to initial state index changing.

- Wen, Severa, Zeng, Luttrell, and Jin, 2002

A general selective encryption approach for fixed and variable length codes (FLC and VLC) is proposed in [11]. FLC and VLC codewords corresponding to important information carrying fields are selected. Then, each codeword in the VLC and FLC (if the FLC code space is not full) table is assigned a fixed length code index, when we want to encrypt the concatenation of some VLC (or FLC) codewords, only the indices are encrypted (using DES). Then the encrypted concatenated indices are mapped back to a different but existing VLC. The encryption process compromises the compression efficiency. Indeed, some short VLC codewords (which are the most probable/frequent) can be replaced by longer ones. This is antagonistic with the entropy coding idea. The proposed scheme is fully compliant to any compression algorithm that uses VLC or FLC entropy coder.

- Zeng and Lei, 2003

In [12], selective encryption in the frequency domain (8×8 DCT and wavelet domains) is proposed. The general scheme consists of selective scrambling of coefficients by using different primitives such as selective bit scrambling, block shuffling, and/or rotation. In wavelet transform case selective bit scrambling and block shuffling is done. In selective bit scrambling the first nonzero magnitude bit and all subsequent zero bits if any give a range for the coefficient value. These bits have low entropy and thus highly compressible and all remaining bits called refinement bits are uncorrelated with the neighbouring coefficients. In this scheme, sign bits and refinement bits are scrambled. In block shuffling, the basic idea is to shuffle the arrangement of coefficients within a block in a way to preserve some spatial correlation; this can achieve sufficient security without compromising compression efficiency. Each subband is split into equal-sized blocks. Within the same subband, block coefficients are shuffled according to a shuffling table generated using a secret key. Since the shuffling is block based, it is expected that most 2D local subband statistics are preserved and compression not greatly impacted.

In DCT transform case, the 8×8 DCT coefficients can be considered as individual local frequency components located at some subband. The block shuffling and sign bits change can be applied on these "subbands." I, B, and P frames are processed in different manners. For I-frames, the image is first split into segments of macroblocks, blocks/macroblocks of a segment can be spatially disjoint and chosen at random spatial positions within the frame. Within each segment, DCT coefficients at the same frequency location are shuffled together. Then, sign bits of AC coefficients and DC coefficients are randomly changed. There may be many intracoded blocks in P- and B-frames. At least DCT coefficients of the same intracoded block in P- or B-frames are shuffled. Sign bits of motion vectors are also scrambled. It is vulnerable to chosen and known plaintext attacks since it is based only on permutations. In addition, replacing the DC coefficients with a fixed value still gives an intelligible version of the image. This algorithm can be part of permutation based encryption.

- Bergeron and Lamy-Bergot, 2005

A syntax compliant encryption algorithm is proposed for H.264/AVC [13]. Encryption is inserted within the encoder. Using the proposed method allows to insert the encryption mechanism inside the video encoder, providing a secure transmission which does not alter the transmission process. The bits “selected for encryption” are chosen with respect to the considered video standard according to the following rule: each of their encrypted configurations gives a non-desynchronized and fully standard compliant bit stream. This can in particular be done by encrypting only parts of the bit stream which have no or a negligible impact in evolution of the decoding process, and whose impact is consequently purely a visual one. About 25% of I-slices and 10–15% of P-slices are encrypted. Since intracoded slices can represent 30–60%, the encryption ratio is expected to be relatively high. The main drawback of this scheme is the lack of cryptographic security. Indeed, the security of the encrypted bit stream does not depend more on the AES cipher. It depends on the size of the compliant codewords. Hence, the diffusion of the AES cipher is reduced to the plaintext space size. In addition, a bias is introduced in the cipher text. This bias depends on the key size and the plaintext space size.

- Lian, Liu, Ren and Wang, 2006

This scheme is proposed for AVC [14]. During AVC encoding, such sensitive data as intra prediction mode, residue data and motion vector are encrypted partially.

Among them, intra prediction mode is encrypted based on exp-golomb entropy coding, the intra macroblocks DCs are encrypted based on context based adaptive

variable length coding, and intra macroblocks ACs and the inter macroblocks MVDs are sign encrypted with a stream cipher followed with variable length coding. The encryption scheme is of high key sensitivity, which means that slight difference in the key causes great differences in cipher video and that makes statistical or differential attack difficult. It is difficult to apply known plaintext attack. In this encryption scheme, each slice is encrypted under the control of a 128 bit sub-key. Thus, for each slice, the

brute force space is 2^{128} ; for the whole video, the brute force space is 2^{256} (the user key is of 256 bit). This brute force space is too large for attackers to break the cryptosystem. According to the encryption scheme proposed here, both the texture information and the motion information are encrypted, which make it difficult to recognize the texture and motion information in the video frames.

3. CONCLUSION

Although an important and rich variety of encryption algorithms have been proposed in literature, most of the algorithms are not secure against cryptanalytic attack. So these algorithms are not suitable for applications which demand high security. It is difficult for a single algorithm to satisfy all performance parameters. We can conclude that it is a challenge for researchers to design an encryption algorithm which satisfies all parameters like visual degradation, speed, encryption ratio, compression friendliness and cryptographic security. Promising future directions of research include more emphasis on key management, resolving the conflict between compression and encryption, and finding ways to change the selection criteria dynamically. Moreover, none of the techniques have used Elliptic curve cryptography, hybrid encryption algorithm

ACKNOWLEDGEMENTS

We give special thanks to Mrs.Uzma, HOD, Department of Computer Science & Engg., RIT Raipur, Chhattisgarh, India, Whose feedback helped me to improve my article.

References

- [1] Borie J., Puech W., and Dumas M., “Crypto-Compression System for Secure Transfer of Medical Images”, *2nd International Conference on Advances in Medical Signal and Information Processing (MEDSIP 2004)*, September 2004
- [2] H. Cheng and X. Li, “Partial Encryption of Compressed Images and Video,” *IEEE Transactions on Signal Processing*, 48(8), 2000, pp. 2439-2451.
- [3] M. Van Droogenbroeck and R. Benedett, “Techniques for a Selective Encryption of

Uncompressed and Compressed Images," Proceedings of Advanced Concepts for Intelligent Vision Systems (ACIVS) 2002, Ghent, Belgium, September 9-11, 2002.

[4] M. Podesser, H.-P. Schmidt and A. Uhl, "Selective Bitplane Encryption for Secure Transmission of Image Data in Mobile Environments," *5th Nordic Signal Processing Symposium, on board Hurtigruten, Norway, October 4-7, 2002.*

[5] A. Pommer and A. Uhl, "Selective Encryption of Waveletpacket Encoded Image Data: Efficiency and Security," *Multimedia Systems*, Vol. 9, No. 3, 2003, pp. 279–287, DOI: 10.1007/s00530-003-0099-y.

[6] J. Meyer and F. Gadegast, "Security Mechanisms for Multimedia Data with the Example MPEG-1 video," Project Description of SEC MPEG, Technical University of Berlin, 1995.

[7] G.A. Spanos and T.B. Maples, "Performance Study of a Selective Encryption Scheme for the Security of Networked Real Time Video," in *Proceedings of the International Conference on Computer Communications and Networks, 1995, pp. 2-10.*

[8] C. Shi and B. Bhargava, "A Fast MPEG Video Encryption Algorithm," in *Proceedings of the 6th ACM International Conference on Multimedia, 1998, pp. 81–88.*

[9] C. Shi, S. Y. Wang, and B. Bhargava, "MPEG Video Encryption in Real-time using Secret Key Cryptography," in *Proceedings of the International Conference on Parallel and Distributed Processing Algorithms and Applications, 1999, pp. 191–201.*

[10] C.-P. Wu and C.-C. J. Kuo, "Fast Encryption Methods for Audiovisual Data Confidentiality," in *Proceedings of SPIE, 2001, Vol. 4209, pp. 284–295.*

[11] J. Wen, M. Severa, W. Zeng, M. H. Luttrell, and W. Jin, "A Format-Compliant Configurable Encryption Framework for Access Control of Video," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 6, 2002, pp. 545–557.

[12] W. Zeng and S. Lei, "Efficient Frequency Domain Selective Scrambling of Digital Video," *IEEE Transactions on Multimedia*, Vol. 5, No. 1, 2003, pp. 118–129.

[13] C. Bergeron and C. Lamy-Bergot, "Compliant Selective Encryption for H.264/AVC Video Streams," in Proceedings of the 7th IEEE Workshop on Multimedia Signal Processing, 2005, pp. 1–4.

[14] Shiguo Lian, Zhongxuan Liu, Zhen Ren and Haila Wang, "Secure Advanced Video Coding Based on Selective Encryption Algorithms," *IEEE Transaction on Consumer Electronics*, Vol. 52, No. 2, 2006, pp. 621-629

Energy Storage Systems in DGR Based Microgrid

K Shravani¹, S.Nagini²

*(EEE, KL University, INDIA)

** (EEE, NOVA College of Engineering, INDIA)

ABSTRACT

After digging deep into the current status of the state of energy storage technologies dominated by batteries in rural India, Energy Storage Systems qualitative analysis, advantages, disadvantages and applications for Microgrid have been discussed in this paper. Various aspects for improving the power quality, continuity, reliability of Microgrid have been suggested by addressing issues like providing ride through, making non-dispatchable power into dispatchable, improving overall performance of power system by developing adequately designed ESS. Storage Systems of importance in this paper are SCES, SMES design, development and deployment. The platform used for presenting theory with proper simulations and result analysis is MATLAB where valid assumptions and specifications of various power system components are constructed using strong basic fundamentals of engineering electric power systems.

Keywords - Microgrid, Interconnected mode, Storage Systems, Super Conducting Energy Storage (SCES), Super conducting Magnetic Energy Storage (SMES), Distributed Generation (DG), Distributed Generation Resources (DGR), Energy Storage Systems (ESS).

I. INTRODUCTION

Vision 2012 has set India on an overarching target of providing electricity to its billion populations by 2012. Power systems market is getting deregulated day by day deploying more localized systems nearer to load centers. By the year 2012, India's peak demand would be 157,107 MW. Based on the demand projections made in the 16th Electric Power Survey, over 1,00,000 MW additional generation

capacity needs to be added by 2012 to bridge the gap between demand and supply of power. But still 40% of the households, mostly in rural areas have no access to electricity even in 2009. At this demanding situation Microgrid is an excellent solution in reaching these whopping targets at fast pace. Microgrid is a collection of renewable decentralized Distributed Generation Resources (DGR) with proper storage technologies and power quality conditioners promising to bridge the gap of supply in rural India. In Microgrid Energy Storage Systems (ESS) play very important role in unifying, distributing and augmenting the

capabilities of alternative and renewable Distributed Generation Systems (DGS). Unlike conventional generation, transmission and distribution systems Microgrid is exposed to load and source fluctuations to a larger extent since the transmission system is absent in its construction. ESS capable of enhancing energy stabilization, ride through capability and dispatch-ability makes up for seasonal variations, reaching to fast transient power quality needs contributing to efficient energy management policies and faster economic investments in Microgrid. In designing and developing ESS a variety of power ratings Kilowatts to Megawatts scale and energy discharge ratings Millisecond to Hour scale are needed to match the wide spectrum of energy storage applications, in the current and forecasted architecture of Microgrid. They can be connected in parallel with the load to supply unscheduled demand permitting DGR to have a smooth steady state. The ride through capability of these systems promises proper amounts of energy to loads in the absence of DGR. Dispatch-ability for certain time regardless of power produced by DGR brings worthiness due to its availability and commit-ability. ESS stabilizes and permits DGR's to run with a constant and stable

output despite fluctuations. ESS along with DGR can be classified as that which stores energy directly in electrical form (Batteries) and the other type which stores energy in various forms finally converting to electrical.

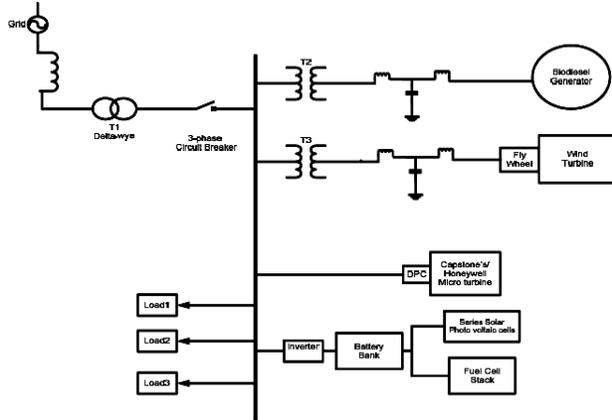


Fig 1: Architecture of Microgrid with a typical Storage System in Interconnected mode.

II. TYPES OF STORAGE SYSTEMS

From application point ESS can be divided into three categories namely Short Term Storage Systems (STS) - maintains energy reserves sufficient to provide rated power from a few milliseconds up to a minute [2]. Faster responses of power conversion systems with storage technologies provide power within fraction of a cycle when the grid supply fails or a voltage depression occurs. Medium Term Storage Systems (MTS) - maintains energy reserves sufficient to provide rated power from few minutes up to few hours. They can be used mainly for peak shaving, network reinforcements, area control, frequency regulation and rapid reserve. They can reduce amount of spinning reserve capacity, resulting in better economic loading and less mechanical and thermal stresses on generators. Long Term Storage Systems (LTS) hold energy for durations ranging from a few hours to weeks or months. The existing distribution systems are greatly suffering from-frequent interruptions due to overloads, insufficient generation, load imbalance, harmonics, transients, sags, swells, The major contributing factors for proper design and implementation of ESS for power distribution systems, both in grid-connected mode as well as in the grid-independent mode/Isolated mode are identification of problems, specifying the energy

density, power density, voltage, current and efficiency requirement, optimizing the specifications in accordance with applications. In power deficient countries like India where scheduled and unscheduled power outages take place and quality of power supply is also poor, areas of exploration includes ESS optimal utilization, dispatch-ability of power by DGR, transient stability improvement and, dynamic stability integrated hybrid systems along with efficiencies.

III. SMES

This system has numerous advantages in electrical power system applications over other conventional means of electrical energy like pumped hydro energy storage, compressed air energy storage etc. Apart from being the most efficient, SMES has attracted the attention because of its fast response in switching over from charging mode to discharging mode, high rate of energy discharge capability. Detailed designs of massive SMES installations (of the order of 1GWh) for load levelling and peak shaving have been evolved. SMES can store electricity and discharge continuously. Super conductivity is a flow of electric current without electrical resistance and is achieved when electrical resistance is made zero by cooling certain chemical compounds. In super conducting state a constant magnetic field is continuously produced even if electricity is conducted, the resultant electrical energy can be stored as magnetic energy. This SMES system uses zero electrical resistance phenomenon to store electricity and achieve an extremely high efficient input and output of electricity. And as a result, the system can be used to provide a stable power supply and to improve quality of power. SMES can be used as a back-up electricity source which can compensate the load damage caused by an instant voltage drop due to lightning, and other factors. It can supply reliable power even in occasions of frequency fluctuations, drastic load changes. Chubu Electric Power Co Inc of Japan with co-operation from Furukawa Electric Co Ltd installed an SMES with 10,000KW of output power.

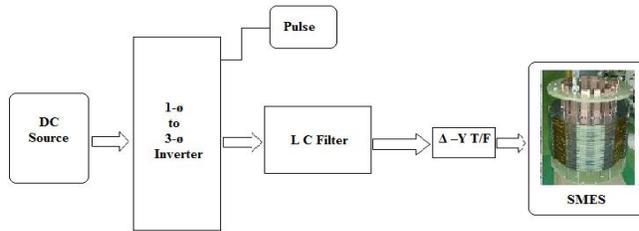


Fig 2: Typical configuration of Single source SMES system.

IV. SMES DESIGN

This system consists of three main parts namely super conducting coil, power conditioning system and cryogenically cooled refrigerator. SMES systems store energy in the magnetic field created by the flow of DC in a superconducting coil which has to be cryogenically cooled to a temperature below its superconducting critical temperature. Once the superconducting coil is charged, current will not decay and the magnetic field can be stored indefinitely. The stored energy can be released back to the network by discharging the coil. The power conditioning system uses a universal bridge to transform Alternating Current to Direct current and Direct Current back to AC. The bridge accounts for 2-3% energy loss in each direction. SMES loses the least amount of electricity in the energy storage process compared to other methods of storing energy. SMES systems are highly efficient with round trip efficiency greater than 95%. With the discovery of high Tc ceramic superconductors the application of SMES in power system may acquire more commercial viability. At present SMES is currently devoted to improve power quality, several 1MW units are used for power quality control in installations around the world especially at manufacturing plants requiring ultra-clean power. If SMES were to be used for utilities it would be a diurnal storage device, charged from base load power at night and meeting peak loads during the day. Cryogenics are a necessity. A robust mechanical structure is usually required to contain the very large Lorentz forces generated by and on the magnet coils. The dominant cost for SMES is the superconductor, followed by the cooling system and the rest of mechanical structure where research is to be carried out in reducing the cost [7]. To achieve commercially useful levels of storage around 1GW-h, a SMES installation would need a loop of 100 miles (160km). Fabrication of bulk cable suitable to carry high currents will be challenging but not impossible.

Until room temperature super conductors are found, the 100 mile loop of wire has to be contained within a vacuum flask of liquid nitrogen.

V. ENERGY CALCULATIONS

The magnetic energy stored by a coil carrying a current is given by one half of the inductance of the coil times the square of the current [7].

$$E = \frac{1}{2} L I^2 \text{ Where}$$

E= energy measured in joules

L= inductance measured in henries

I= current measured in amperes

If we consider cylindrical coil with conductors of a rectangular cross section, mean radius 'R', 'a' and 'b' respectively width and depth of the conductor, 'f' the form function different for different shapes of the coil, $\hat{E}(xi)$ and $\Delta(\delta)$ being the two parameters that characterize the dimensions of the coil. Now with above assumptions the magnetic energy stored in such a cylindrical coil can be written as

$$E = \frac{1}{2} F(\hat{E}, \Delta) . R . N^2 . I^2$$

Where E= energy measured in joules

I= current measured in amperes

F(\hat{E}, Δ)= form function, joules per ampere-meter.

N= number of turns of coil.

VI. SCES

Super capacitors are less weighty than that of battery of same energy storage capacity, a fast access to the stored energy is the hallmark of super capacitors [1]-[8]. Basic needs like super fast charging and improved discharging cycles up to 10^6 times the charging times is possible with super capacitors. High storage capacities independent of number of charging and discharging cycles, energy densities of the order of 10 to 100 times to that of traditional capacitors, life cycles of 25-30 years, high efficiencies, higher power densities greater than batteries are essential features of super capacitors.

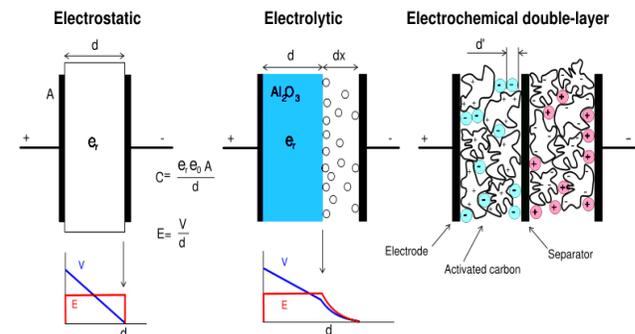


Fig3. Types of Super capacitor/ultra capacitors

Typical characteristics of Super-capacitors include nominal voltages ranging from 2.3 to 400 volts, rated current ranging from 3-600 amperes, operating temperatures of -40°C to 85°C, maintenance free and very low leakage currents. Constructional features of Super-capacitors include double layers. The energy is stored by charge transfer at the boundary between electrode and electrolyte. The amount of stored energy is function of the available electrode and electrolyte surface, size of the ions and level of electrolyte decomposition. The two electrodes made of activated carbon provide a high surface area part, defining so energy density of the component on the electrodes, current collectors with a high conducting part assure the interface between the electrodes and the connections of the super capacitor.



Fig4: Maxwell technologies Super-capacitors (“MC” and “BC” series up to 3000 farads).

VII. SCES WITH FUEL CELL

The Fuel Cell Stack block in Mat lab implements a generic model parameterized to represent most popular types of fuel cell stack fed with hydrogen and air. PEMFC-50KW-625Vdc block is utilised here to transform assumed theory into simulative result.

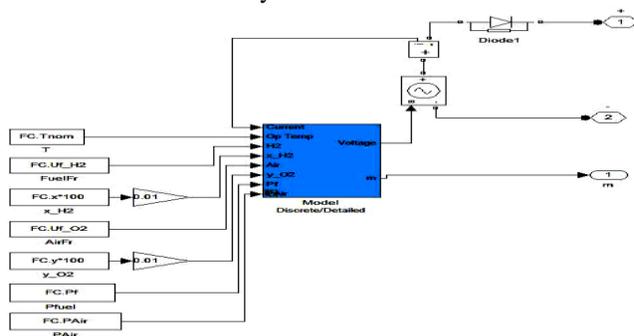


Fig5: A PEMFC-50KW-625Vdc Fuel Cell.

A diode is used to prevent the negative current into the stack. Nominal voltages, open circuit voltage, operating points have to be calculated with the basic equations governing the output from the fuel cell module.

$$E^{OC} = N [E^n - A \ln(i^o)] \text{ and } A = RT/Z\alpha F$$

Where $R = 8.314 \text{ J/(Mol K)}$
 $F = 96485 \text{ A}$
 $Z = \text{Number of moving electrons}$
 $E^n = \text{Nernst voltage}$
 $I^o = \text{Exchange current}$

An IGBT based inverter is used to convert the DC from fuel cell to AC and harmonic filter is used to reduce the harmonics. A Super Capacitor bank is used for storing the output and also for peak shaving. Power quality conditioning is done by proper synchronization of inverter, filter and super capacitor bank. A PWM Generator provides pulses to the inverter at appropriate time intervals playing a major role in the quality.

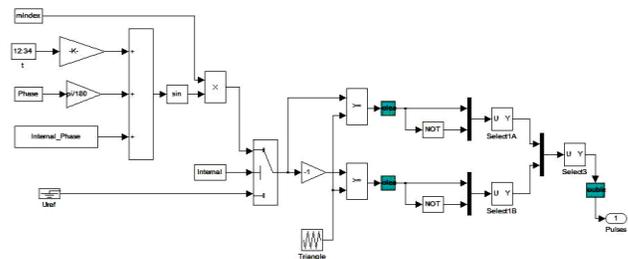


Fig6: PWM generator for IGBT Inverter.

VIII. SCES WITH 3-DGR MICROGRID

A Microturbine generation system along with wind power generation system and fuel cell power generation system are integrated with appropriate combinational outputs guided by PID controllers for wind system to get assumed torque, speed, and pitch angle at variable wind speeds. The configuration of PID is shown below.

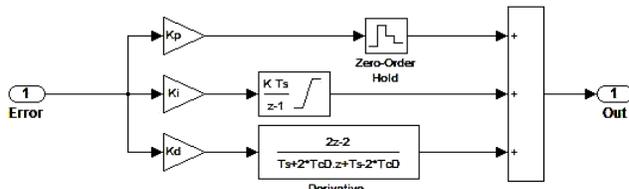


Fig 7: PID Controller for Wind Gen System.

The output from the High Speed Gas based Micro turbine is heavily dependent on the configuration of steam turbine and governor. The speed governing system consists of a proportional regulator, a speed relay, and a servomotor controlling the gate opening. The steam turbine has four stages, each modeled by a first-order transfer function. The first stage represents the steam chest while the three other stages represent either Reheaters or crossover piping. The boiler is not modeled and boiler pressure is constant at 1.0 PU.

IX. RESULT ANALYSIS

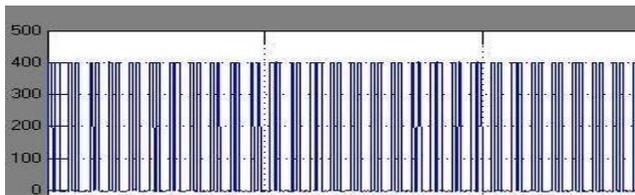


Fig 8: V_{ab} of inverter for SMES system with 400 V as input.

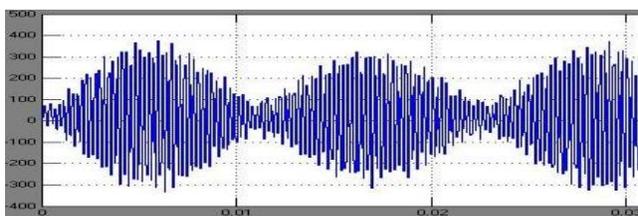


Fig 9: V_{ab} of Load for SMES system with 400V as input.

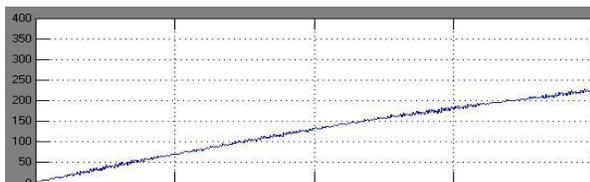


Fig 10: Single Source SMES system Phase Currents (I_{ab}~400 amps).

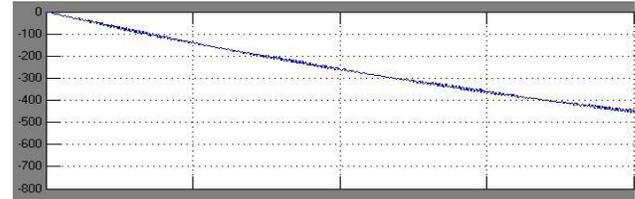


Fig 11: I_{bc}~800 amps in 0.1 seconds.

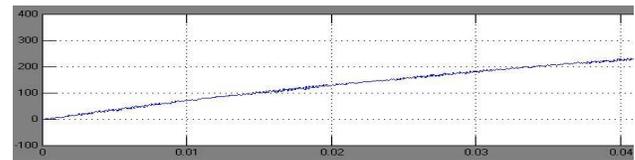


Fig 12: I_{ca}~400 amps in 0.1 seconds.

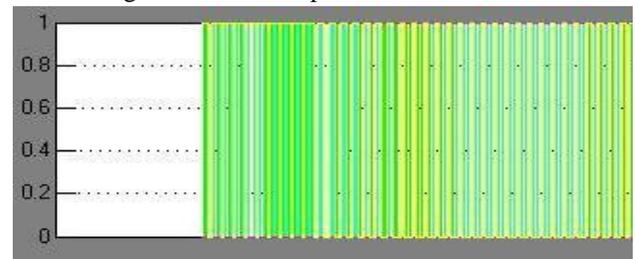


Fig 13: Fuel Cell (PEMFC-625V-50KW) based SCES system pulses.

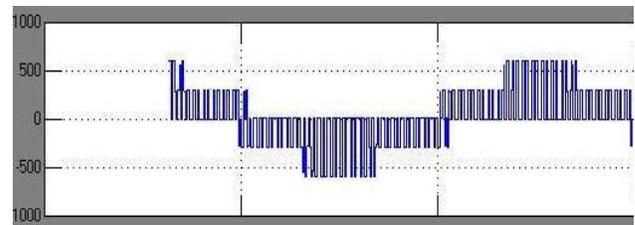


Fig14: Fuel cell based SCES system V_{ab}.

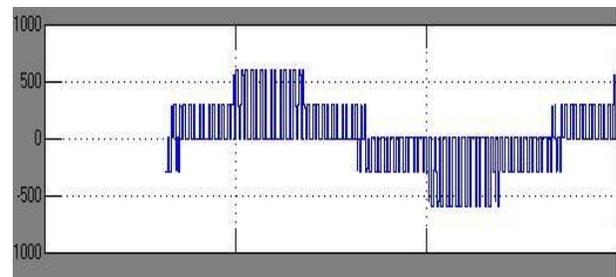


Fig 15: Fuel cell based SCES system V_{bc} .

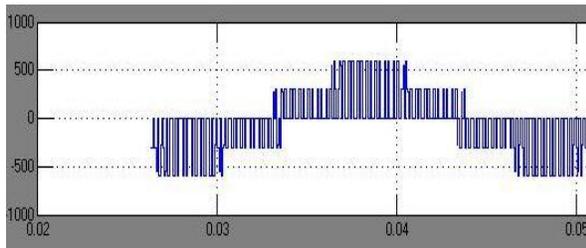


Fig 15: Fuel cell based SCES system V_{ca} .

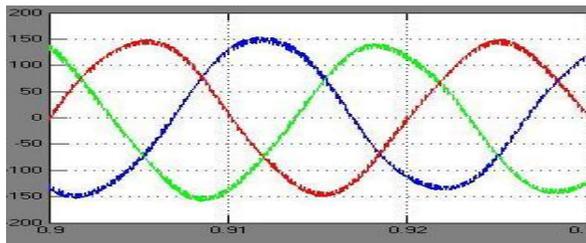


Fig 16: 3-DGR Microgrid with SCES system
 $V_{abc} \sim 150$ Volts.

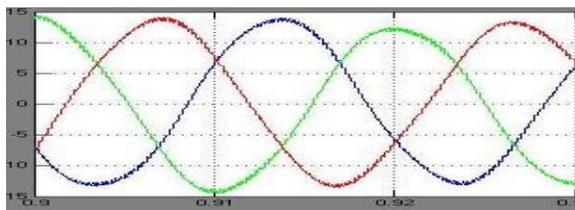


Fig 17: 3-DGR Microgrid with SCES system $I_{abc} \sim 15$ amps.

IX. CONCLUSION

Distributed generation with load and source at vicinity to each other is on fast track in India with lot of importance given in the current financial year. At this peak these kind of unconventional storage systems deployment will definitely prove successful in maintaining the quality and quantity on par with demand. SMES need to be supported with further research in reducing the cost of superconductor and cryogenic facilitation. SCES along with power electronic conditioning systems will be a sure success in many industrial and renewable generation systems design, development and deployment. The major benefits that can be achieved are distribution upgrade deferral, transmission upgrade deferral, availability-

based use, end user electric service reliability. This would also lead to positive impact on environment by virtue of reduced fuel consumption, reduced emissions of dangerous green house gases, and reduced global warming and renewable capacity optimal utilization.

REFERENCES

- [1] Kuldeep Sahay, Bharti Dwivedi, "Super capacitors energy storage system for power quality improvement" in Journal of Electrical Systems x-x (2009).
- [2] W. Kramer, S. Chakra borty, B. Kroposki, and H. Thomas "Technical Report on Advanced Power Electronic Interfaces for Distributed Energy Systems Part 1, Systems and Topologies" NREL/TP-581-42672 -March 2008.
- [3] Paul Denholm, Erik Ela, Brendan Kirby, and Michael Milligan, "The Role of Energy Storage with Renewable Electricity Generation Technical Report "NREL/TP-6A2-47187, January 2010.
- [4] PIER FINAL PROJECT REPORT, "An assessment of battery and Hydrogen energy storage systems Integrated with wind energy Resources in California" September 2005- CEC-500-2005-136.
- [5] Kuldeep Sahay, Bharti Dwivedi, "Energy storage technology for performance enhancement of power system" in Electrical Power Quality & Utilization Magazine, Volume 4, Issue 1.
- [6] Robert B. Schainker, IEEE, "Executive Overview- Energy Storage options for a Sustainable Energy Future".
- [7] K.H.J.Buschow, "Concise nyclopaedia of Magnetic & superconducting materials- Second edition", ELSEVIER.
- [8] Kuldeep Sahay and Bharti Dwivedi, "Design and Analysis of Supercapacitors Energy Storage System for Energy Stabilization of Distribution Network", Department of Electrical Engineering, Institute of Engineering & Technology, Uttar Pradesh Technical University, Lucknow, UP 226021.
- [9] Phatiphat Thounthong, Member, IEEE, St'ephane Ra'el, and Bernard Davat, Member, "Analysis of Supercapacitor as Second Source Based on Fuel Cell Power Generation", IEEE- IEEE

Transactions on Energy Conversion, vol. 24, no. 1, March 2009.

- [10] Guest Editorial, "Special Section on Vehicular Energy-Storage Systems", IEEE Transactions on Vehicular Technology, vol. 58, no. 8, October 2009.
- [11] H. Maker H. Gualous R. Outbib Monica E. Romero, "Control of Bidirectional DC-DC Converter for Supercapacitor Automotive Application", January 22, 2008.
- [12] BENT SORENSON, "Renewable Energy Conversion, Transmission and Storage".

Optimization of process parameters in Abrasive Flow Machining with Particle Swarm Optimization

Jose Cherian

Research Scholar,,Dept of Mechanical Engineering,
Karpagam University
Coimbatore, India

Dr Jeoju M Issac

Associate Professor, Dept of Mechanical Engineering,
MA College of Engineering
Kothamangalam

Abstract—Abrasive flow machining (AFM) process is a non-traditional finishing process used for polishing and radius difficult to reach surfaces by the abrading action of the abrasives. The material to be machined is taken in the form of a cylinder. The abrasives are taken in the work piece and rotated at high RPM.AFM can be used to produce high surface finish.Vaiious process parameters are abrasive size, Machining time, Hardness of abrasives and speed of abrasives. The experimental results reveals that the efficiency of the process strongly linked to the mechanical properties of the machined material and machining time. This technique offers good surface finish without affecting closest geometrical tolerances of materials.

Keywords-Abrasive flow machining, Surface finish, Particle swarm optimization; fitness function.

I. Introduction

AFM is widely used as a finishing process to finish complicated shapes and profiles. This technique developed in 1960s.AFM is mainly classified into three types. (1) one way AFM (2) two way AFM (3) Orbital AFM.The polymer abrasive medium which is used in this process, posses easy flow ability, better self deformability and fine abrading capability. The ability of media in AFM process to finish difficult to reach areas, to follow complex contours and to simultaneously work on multiple edges and surfaces, makes it more versatile than other finishing process. A thickness of 1 to 10 μ m can be removed by this process.AFM reduces surface roughness by 75 to 90 percent on cast and machined surfaces. It can be used to produce uniform, repeatable and predictable results on an impressive range of polishing operations. Important feature which differentiates AFM from other finishing process is that it

is possible to control and select the intensity and location of abrasion through fixture design, medium selection and process parameters. Optimization of process parameters in AFM with neural networking has been done by R.K.Jain and V.K.Jain[1] The objective of the present paper is to optimize the process parameters of the process through Particle Swarm Optimization method. –

2. Particle Swarm optimization (pso)

Particle swarm Optimization is a stochastic optimization technique developed by Dr Eberhart and Dr Kennedy, inspired by social behavior of bird flocking or fish schooling.PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations. In every iteration each particle is updated by following two best values. The first one is the best solution (fitness) it has achieved so far. This value is called pbest. Another best value that is tracked by the particle swarm optimizer is the best value obtained so far by any particle in the population. This value is the global best and called gbest. When a particle takes part of the population as its topological neighbors, the best value is called local best or lbest. After finding the two best values, the particle updates its velocity and positions.PSO is easy to implement compared to GA. This method is effectively applied in the areas of function optimization, artificial neural networking and fuzzy system control.

2.1 PSO algorithm

Each particle keeps track of its coordinates in the problem space, which are associated with the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called pbest. Another best value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the neighbors of the particle; the best value is a

global best and is called gbest[2,3]. The steps involved in PSO are outlined as follows.

Phase I - Initialization

The particle parameters are randomly generated in solution space. This provides a set of values to begin the iteration.

Phase II – Evaluation

- (1) The *fitness* of the particle is evaluated.
- (2) In every iteration, each particle is updated by following two "best" values.
 - (i) Best solution (fitness) it has achieved so far (pbest).
 - (ii) The best value, obtained so far by any particle in the population (gbest).
- (3) After finding the two best values, the particle updates parameters.

Phase III – Stopping Criteria

The iterative procedure is stopped if one of the following criterions is met:

- (1). Maximum change in best fitness smaller than specified tolerance.
- (2). If maximum number iterations are attained.

2.2 Pseudo code for the algorithm

```

    For each particle
    {
        Initialize particle parameters
    }
    END
    Do
        {For each particle
        Calculate fitness value (If the fitness value is better than the
        best fitness value (pBest) in history set current value as the
        new pBest)
        }
        END
        (Choose the particle with the best fitness value of all the
        particles as the gBest)
        For each particle
        {
            Update particle parameters
        }
        END
    (While maximum iterations or minimum error criteria is
    attained)
    
```

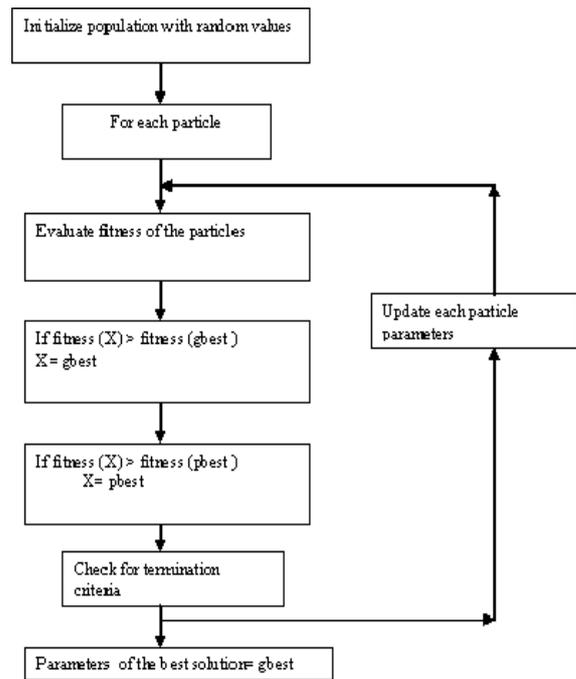


Figure 1. Pso flow chart

3 Experimentation

Machining experiments were carried out with aluminum as the work piece material. The medium was silicon carbide and putty mixture. Experiments were carried out by changing rpm of the work piece. Percentage concentration is defined as the ratio of weight of the abrasives and total weight of the medium multiplied by 100. When the rpm changes the linear velocity of abrasive changes from 40 to 90cm/min. The variation of and surface finish with respect to velocity, abrasive size were measured. The data collected from these experiments were used for PSO analysis.

3.1 Experimental details

PROCESS PARAMETERS

- Abrasive flow speed (v) cm/min
 - Vmin < v < vmax
- Percentage concentration of abrasives (c)
 - C min < C < C max
- Abrasive mesh size
 - d min < d < d max
- Number of cycles (n)

3.2 Constraint

The constraint evaluated is Surface roughness constraint (Ra). It indicates the quality of machined surface (surf test equipment is used to measure surface roughness).

3.3 Work piece material

The work piece material used in the study was Aluminium. They were in the form of cylindrical tubes of diameter 4cm and length 10cm.

3.4 Abrasive used

Silicon Carbide(SiC)

Table 1. Experimental results

3.5 Results obtained

The analysis was made by running PSO program(C language), for 50 iterations. The fitness function is evaluated using following relation

$$\text{Fitness function} = \sum_{i=1}^n (Ra(i) - R_{ra}(i))^2$$

Empirical model for surface roughness

$$R_a = K v^x c^y d^z n^p$$

The results obtained are

Value of K=28270

$$X = - 1.82$$

$$Y = - 1.32$$

$$Z = - 1.40$$

$$P = - 0.23$$

Empirical model by PSO

$$R_a = 28270 v^{-1.82} c^{-1.32} d^{-1.40} n^{-0.2258}$$

4 Conclusions

For solving machining optimisation problems, various conventional techniques have been used so far, but they are robust and have problems when they are applied to AFM, which involves a number of variables and constraints. They are non-linear also. To overcome the above problems, particle swarm optimisation is used in this work. Particle swarm optimisation converges to the global optimal solution faster. *The PSO technique was found to converge to optimum in a faster rate. PSO is a generalised technique and can be easily modified. The method requires only primitive mathematical operations, so it is*

computationally inexpensive in terms of memory and requirements and speed.

Comparison of PSO results with those of neural networking shows that the PSO algorithm is more effective for optimization of machining parameters. The algorithm is simpler in PSO. Number of iterations required to reach optimum value is also less in PSO.

| S.I no | Abrasive flow speed (cm/min) | Concentration (c) | Mesh size | n (NO) | Hardness | Roughness Ra (µm) |
|--------|------------------------------|-------------------|-----------|--------|----------|-------------------|
| 1 | 40.60 | 45.03 | 100 | 50 | 160 | 1.52 |
| 2 | 51.50 | 37.0 | 100 | 20 | 160 | 1.82 |
| 3 | 51.50 | 45.0 | 150 | 50 | 160 | 1.40 |
| 4 | 65.20 | 45.0 | 100 | 50 | 170 | 0.95 |
| 5 | 81.20 | 45.0 | 100 | 50 | 170 | 1.60 |

References

- [1].R.K.Jain and V.K.Jain Optimum selection of machining conditions in abrasive flow machining using neural network—journal of Materials processing Technology 108 (2000) 62- 67
- [2] Wen Jye Shyr “An optimum design course supported by the particle swarm optimization algorithm for undergraduate students” World transactions on engineering and technology Education Vol.5 pp 183-186, 2006.
- [3].LIU Yi jian,ZHANG jian ming, WANG Shu-quig “Parameter estimation of cutting tool temperature nonlinear model using PSO algorithm” journal of Zhejiang University Science pp 1026-1029, 2005.
- [4]Sijo M.T and Biju N. “Taguchi Method for Optimization of cutting parameters in turning operations” Proc of International conf on Advances in Mechanical Engineering
- [5]. T. G.Ansalam Raj and V.N .Narayanan Namboothiri “An improved genetic algorithm for the prediction of surface finish in dry turning of SS 420 “ materials manufacturing technology today 47,pp 313-324, 2010.
- [6].J S Senthilkumar ,R Saravanan and P Asokan “ multiple optimization for selection of machining parameters of inconel-718 material turning process” manufacturing technology today pp 3-7, Feb 2008.
- [7]. M Barletta “Progress in Abrasive fluidized bed Machining” Journal of material processing technology(2009).

The Nutrient Content of Organic Liquid Fertilizers in Zimbabwe

Simbarashe Govere^{*1}, Benard Madziwa², Precious Mahlatini³

^{*1,3} (Department of Environmental Sciences & Technology, Chinhoyi University of Technology, Zimbabwe)
² (Zimbabwe Organic Agricultural Producers and Promoters Association, Harare, Zimbabwe)

ABSTRACT

This paper assessed the nutrient content of three organic liquid manures made from Water Hyacinth (*Eichhornia Crassipes*), Russian Comfrey (*Symphytum officinale*) and Pig Weed red-root (*Amaranthus retroflexus*) plants. The liquid manures were made by shredding plant materials and fermenting them in water for 30 days. Samples were analyzed weekly for nitrogen, phosphate and potassium (NPK) and trace elements. Water Hyacinth liquid manure had significantly high N (3.72%) and P (2.86%) contents indicating its suitability as a macronutrient fertilizer. All liquid manures had high K contents, particularly Russian Comfrey (3.90%), hinting against direct foliar application. Pig weed had high levels of Ca, Zn and Mg suggesting its suitability as a sufficient micronutrient fertilizer. All liquid manures had NPK contents greater than common solid organic fertilizers such as cattle manure used in Zimbabwe.

Keywords – Organic farming, liquid manure, Water Hyacinth, Russian Comfrey, Pig Weed

1. Introduction

Organic farming is an internationally regulated, legally enforced and standardized alternative agricultural paradigm that relies on ecological processes, biodiversity and cycles adapted to local conditions with the aim of sustaining the health of soils, ecosystems and people [1]. Originating in

Germany in the early 19th century from the pseudo-scientific roots of biodynamic farming, organic farming has evolved into an ecologically friendly substitute of conventional farming systems which have been shown to have adverse environmental and health impacts due to intensive use of synthetic inputs such as pesticides, herbicides and fertilizers [2, 3, 4].

The substitution of synthetically manufactured agricultural inputs, such as chemical fertilizers, by minimally processed naturally-occurring organic inputs, such as organic fertilizers, forms the core tenet of organic farming [5]. Organic fertilizers are soil amendments containing the minimum contents of, at the minimum, nitrogen, phosphate and potash (NPK) that is derived solely from the vestiges or derivatives of an organism [6, 7]. The NPK in organic fertilizers must come solely from organic reserves that are inherently high in NPK and not from petroleum or ammonia- derived fertilizers like ammonium nitrate or synthetic urea [7]. The application of organic fertilizers has been shown to have positive impacts on soil fertility, soil physical properties and consequently crop yield [8, 9, 10].

Depending on the raw materials, method of production and environmental conditions organic fertilizers vary widely in

form and nutrient content . Organic fertilizers include solid organic fertilizers (farm yard manure, green manure and compost) as well as liquid organic fertilizers (plant extracts, compost watery extracts, compost leachate, compost teas, liquid manures and manure teas) [11] .Globally there is prevalent use and

scientific knowledge of solid organic fertilizers. Table 1 below highlights some common solid organic fertilizers used in Zimbabwe and their nutritive capacity. Cattle manure and green manure are the most nutritive organic fertilizers with an average nitrogen (N) content of 1.50% [12]. Elsewhere, beyond Zimbabwe, bat-guano (feces) harvested from caves has been shown to be an excellent source of nutrients with an NPK value of 10-3-1 [7]. Other organic fertilizers with high NPK values include fish meal (10-6-2); blood meal (12-0-0) and bone meal (3-15-0) [13].

Table 1: Mean nutrient quality of solid organic fertilizers used by small farmers in Zimbabwe

| Fertilizer type | N (%) | P (%) | K (%) |
|-----------------|-------|-------|-------|
| Cattle Manure | 1.50 | 0.15 | 0.78 |
| Leaf litter | 1.40 | - | - |
| Anthill soil | 0.23 | 0.05 | - |
| Compost | 0.34 | 0.12 | - |
| Crop residue | 0.45 | 0.06 | - |
| Legumes | 1.50 | 0.08 | - |

Source: [12]

Liquid organic fertilizers have largely remained in the background of mainstream scientific literature and what little knowledge exists about them is mainly confined to biodynamic farming literature [14]. Seaweed liquid fertilizer (also known as SLF, kelp or seaweed extract) is a popular liquid fertilizer in India, Europe and America with a history of use dating back from Ancient China [15]. Seaweed extract has been found to contain negligible amounts of N and P but high levels of all trace elements and plant hormones (IAA and IBA) cytokinins, gibberellins and vitamins [16].

In India the traditional liquid fertilizer called Panchagavya, a manure tea made by fermenting cow dung in water, has been shown to have a modest NPK content of 0.03-0.02-0.04 but a high iron content of 0.84%. Other Indian liquid manure teas such as Jeevamrut and Beejamrut are reportedly used not as sources of nutrients but of plant growth hormones [17, 18].

Early biodynamic teachings by Rudolf Steiner in 1924 popularized liquid manures as a common gardening experience in Europe and America [19]. Liquid manures are made by fermenting plant material and extracts in water [20]. The upsurge in

organic farming popularity in Europe and America has turned organic liquid fertilizer production into a big commercial business with products such as Biolizer XN and Tomarite sold on the international market. However their unsubstantiated high N content claims (above 6%) have introduced a lot of controversy with some products being withdrawn following objections from organic certifiers and growers [6].

Typical of Africa, Zimbabwe has traditional agricultural practices which are compatible with organic farming [21]. Rural communities and smallholder farmers generally lack capital to buy synthetic pesticides and inorganic fertilizers and rely heavily on organic fertilizers [22]. In recent years several non-governmental organizations promoting organic farming in Zimbabwe such as the Zimbabwe Organic Agricultural Producers and Promoters Association (ZOAPPA) have emerged. Their work is hampered by limited data on aspects such as organic fertilizers and pesticides.

This research sought to determine the NPK and micronutrient (Ca, Mg, Zn, S) content of liquid manures made from three common plant materials in Zimbabwe; Water Hyacinth (*Eichhornia Crassipes*), Russian Comfrey (*Symphytum officinale*) and Pig Weed red-root (*Amaranthus retroflexus*).

2. Materials and Methods

2.1 Study Site

The study was conducted in Harare (lon: 31.0⁰E and lat: 17 .8⁰S) in Zimbabwe at Tese Tigute Organic Plot in the North –Eastern suburb of Highlands. The study area is an open door learning centre, with an agricultural area characterized by a smallholder organically certified intensive farming system with vegetation marked by the main crop, maize (the staple food grain), vegetables and diversified herbal plants. The area experiences three main seasons, a warm wet season from November to March/April; a cool dry season from May to August and a hot dry season in September/October. The area has a humid sub tropical climate of average temperature 17.95 °C.

2.2 Experimental Design

An experimental design approach was used whereby the 3 treatments of Water Hyacinth (*Eichhornia Crassipes*), Russian Comfrey (*Symphytum officinale*) and Pig Weed (*Amaranthus retroflexus*) fermenting in water acted as the experiments. For each treatment they were 4 replicates.

2.3 Selection of Plant Material

The rationale behind the selection of the plants used for the liquid manure production was based on the ease of availability of the plant, documented plant NPK content and its biomass production. Russian Comfrey (*Symphytum officinale*) is a prolific exotic plant that propagates rapidly by seeds and root stems and produces 4-5 kg of leaves per plant per cutting during the growing season [23]. It has a high nutrient content due to its deep root system that mines minerals from the subsoil. Water hyacinth (*Eichhornia Crassipes*) is a perennial aquatic weed which has proved to be an environmental nuisance in polluted water bodies in Zimbabwe such as Lake Chivero and Lake Kariba [24]. It has a high mineral nutrient absorption and accumulation capacity and thus its application in wastewater treatment, animal fodder and organic fertilizer [25]. Pig Weed red-root (*Amaranthus retroflexus*), apart from being a noxious weed in maize and bean plots, is common along Zimbabwean roadsides and open waste sites. Pigweed red-root belongs to the Amaranth family, a grain crop, and under certain conditions accumulates nitrates in its leaves and stems [26]. The only effective cultural way of managing both Water Hyacinth and Redroot Pigweed the weeds is through cultivation, physical removal and dumping of the plant residue. Given the high biomass produced by the weeds the use of Water Hyacinth and Pigweed in liquid manures could be an ecologically sound way of depositing the weed residue.

2.4 Fermenting Process

In this study a plant to water weight ratio of 1:8 was used to allow the extraction of the highest amount of nutrients. Normal plant to water weight ratios used in both biodynamic and organic farming range between 1:8 to 1:10 [23, 27].

Healthy Comfrey (*Symphytum*) and Amaranthus specimens not yet in the flowering stage were sourced from Tigute Organic Plot. Water Hyacinth

(*Eichhornia Crassipes*) was sourced from Lake Chivero. The plants were collected and cleaned and the root section was cut and discarded. Only the leaves and stem were used in the experiment. Using an analytical balance 5 kg of the plant material was weighed, shredded and then placed in a 40 litre container. Forty liters of water was added and the container was sealed to exclude rainwater, prevent evaporation and contain flies and odors. The material was fermented for 30 days but periodically stirred every 7 days.

2.5 Chemical Tests

Samples of 200ml from each fermenting container were collected, filtered and centrifuged at 10,000rpm for 30 minutes at 4EC. The resulting supernatant was taken for nutrient analysis (%N, % P, %K, %S, %Mg, %Ca, %Zn) by standard chemical analysis as shown in Table 2. Samples were collected after every five days from the 15th day to the 30th day of the study.

Table 2: Chemical Analysis

| Method | Analysis |
|--|-------------------------------------|
| pH | pH meter |
| Nitrogen | Kjeldahl Method |
| Phosphorus | UV-VIS Spectrophotometer |
| Potassium, Zinc, Magnesium, Calcium, Sulphur | Atomic Absorption Spectrophotometer |

The data was analyzed using Genstat (version 6) statistical package.

3. Results

A summary of the results for macronutrients and micronutrients content in liquid manures after 30 days of fermenting are presented in Table 3 below.

Table 3: Mean Macro and micronutrient content in Liquid Manures.

| | Water Hyacinth | Pig weed | Russian Comfrey | Water |
|-----|----------------|-----------|-----------------|-------|
| pH | 6.7 | 6.8 | 7.8 | 6.8 |
| TDS | 980±40 | 1844±20 | 972±88 | 91 |
| %N | 3.72±0.33 | 1.54±0.37 | 2.90±0.1 | 0.00 |
| %P | 2.86±0.41 | 2.98±0.24 | 2.94±0.05 | 0.00 |
| %K | 2.89±0.02 | 2.01±0.4 | 3.90±0.06 | 0.00 |
| %S | 0.91±0.21 | 0.70±0.34 | 1.60±0.34 | 0.05 |
| %Mg | 0.08±0.02 | 0.16±0.03 | 0.08±0.02 | 0.00 |
| %Ca | 0.06±0.01 | 0.38±0.11 | 0.06±0.03 | 0.00 |
| %Zn | 0.05±0.01 | 0.05±0.01 | 0.04±0.02 | 0.05 |

3.1 Nitrogen

Results show that Water Hyacinth liquid manure had the highest N content (3.72%) after 30 days of fermentation (see Fig 1). Russian Comfrey liquid manure had an N content of 2.90% and Pig Weed liquid manure had the lowest N content of 1.54%. There was a significant difference (p= 0.05) between the N content in all the liquid manures (see Fig 1). All the liquid manures had an N content higher than most conventional organic fertilizers (compare Table 1). This suggests that liquid manures can be employed as more nutritive substitutes to solid organic fertilizers such as legume green manure and cattle manure.

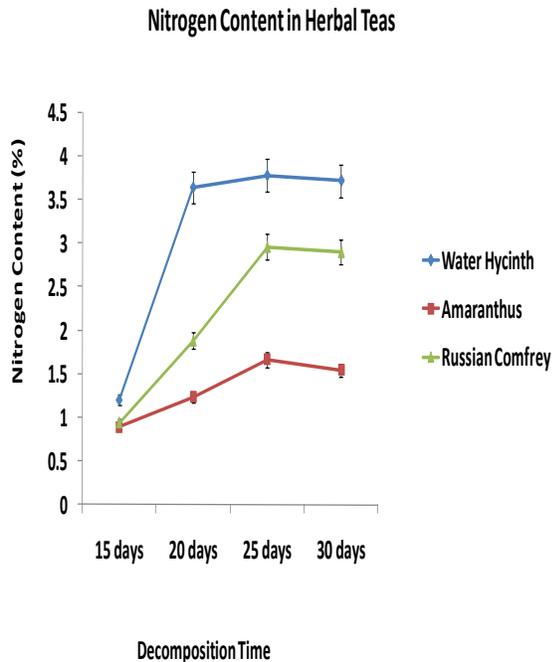


Figure 1: N content in liquid manures

The source of the nitrogen in the liquid manures is the inherent N contained in the plant matter. Water Hyacinth is known to have high protein content [28]. In organic agriculture 6.2% protein content is equivalent to 1% nitrogen content [6]. Water Hyacinth leaves may contain as high as 15.17% protein which might yield more than 2% nitrogen during decomposition [29, 30]. There is great variation in the documented total nitrogen content of Water hyacinth leaves as reported by various authors depending on whether it is collected from a polluted (nutrient rich) or an unpolluted (nutrient poor) water source [25]. The Water Hyacinth for this study was collected from Lake Chivero a polluted water source and thus might contain high N content.

It took 20 days for Water Hyacinth liquid manure to reach its maximum N content compared to the other two liquid manures which took 25 days to ferment indicating a faster decomposing rate for the Water Hyacinth plant. Russian Comfrey liquid manure and Pig Weed liquid manure had a slight reduction in the N content by the 30th day suggesting N losses due to volatilization can occur.

The faster decomposition time for the Water hyacinth plant and the N losses via volatilization for the Russian Comfrey and Pig Weed can be explained by the C: N ratio. Water Hycinth has a C: N ratio of 36:1 [30], Pig Weed 11:1 [31] and Russian Comfrey **9.8:1** [28]. For optimum decomposition to occur microorganisms require a suitable C: N ratio of 30: 1 (between 20/1 and 35/1) [32]. Thus Water Hyacinth has a suitably high C:N ratio which facilitates faster decomposition by microbes. A low C: N ratio results in N losses due to volatilization. Nitrogen loss due to excess nitrogen in the pile (a low C/N ratio) can be over 60%. At a C/N ratio of 30 or 35 to 1, only one half of one percent of the nitrogen will be lost [32].

3.2 Phosphate

Although results show that Pig Weed liquid manure has the highest P content of (2.98%) followed by Russian Comfrey liquid manure (2.94%) and Water Hyacinth liquid manure (2.86%) there was no statistical difference (p= 0.05) between the P content in the liquid manures (see Fig 2). All liquid fertilizers had P content greater than most common organic fertilizers used by small scale farmers in Zimbabwe (see Table 1).

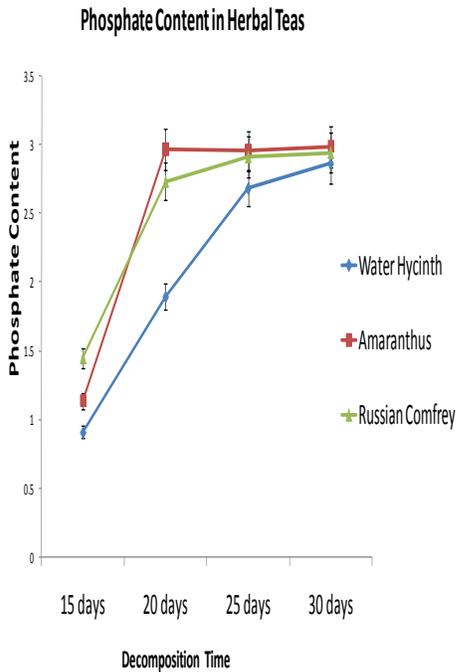


Figure 2: P content in liquid manures. Phosphate in solution exists as $H_2PO_4^-$ and HPO_4^{2-} with the former more abundant at acidic pH ($pH < 7.2$). Plant uptake of $H_2PO_4^-$ is more rapid than HPO_4^{2-} . Comparing the pH values of the liquid manures it can be theorized that Water Hyacinth liquid manure might be the better source of P due to the acidic pH of 6.7 [37]. Phosphate availability is further compromised by Ca^{2+} which precipitates insoluble Dicalcium Phosphate and Hydroxyapatite. Thus the high Ca content in Pig weed liquid manure may inhibit P availability to plants [33].

3.3 Potash

Russian Comfrey liquid manure had the highest potash content of 3.90% whilst Water hyacinth liquid manure had 2.89% and Pig weed liquid manure had 2.01%. The potassium content in all 3 liquid manures was statistically different ($p=0.05$) (see Fig 3). The potassium content in all liquid manures is greater than that of cattle manure (see Table 1). The high K concentration of the liquid manures for seems to discourage foliar application. Nutrient concentration of generally less than 1-2% is employed to avoid injury to foliage [34].

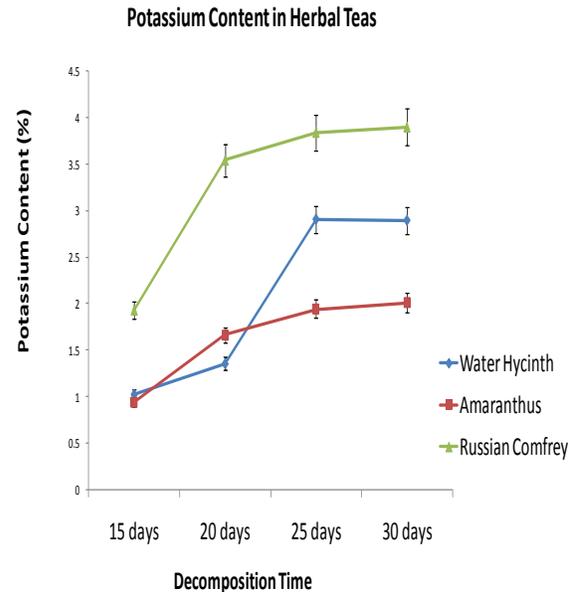


Fig 3: K content in liquid manures.

Effectiveness of solution K for crop uptake is influenced by Ca and Mg ion presence. High concentration of both Ca and Mg inhibit crop uptake of k and is qualitatively assessed by the activity ratio (AR_e^k) equation [33].

$$AR_e^k = a_k / \sqrt{(a_{Ca} + a_{Mg})} \quad (i)$$

AR_e^k is governed by the potassium activity (a_k), calcium activity (a_{Ca}) and magnesium activity (a_{Mg}). The higher the AR_e^k the higher the plant K. Calculating the AR_e^k of the liquid manures shows that Russian Comfrey liquid manure has more available plant K (11.47) compared to Water Hyacinth liquid manure (7.7) and Pig Weed liquid manure (2.73).

3.4 Trace Elements

Results of trace element analysis shows that Russian Comfrey Liquid Manure has the highest sulphur content of 1.6%; Pigweed had the highest Calcium (0.38%) and Magnesium (0.16%) content. Zinc contents were equally high in Water Hyacinth and Pig weed liquid manures. From the results Pigweed might be the best source of all trace elements. The Pigweed Liquid Manure provides the highest amount of Zn, Ca and Mg. Although Pigweed had the lowest Sulphur content (0.7%) it is still sufficient to meet plant requirements.

4. Conclusion

The paper reported on a study that evaluated the nutritional component of liquid fertilizers derived from three common plants in Zimbabwe namely, Water Hyacinth, Russian Comfrey and Pigweed. Water Hyacinth liquid manure had the highest N and P content and on that basis can be concluded as the more superior macronutrient fertilizer. In addition it has no N losses due to volatilization and generally takes less time (25 days) to prepare due to a higher C:N ratio. Generally all liquid manures were of higher nutrient content than common solid organic fertilizers used by farmers.

All the liquid manures had an equally high P content (> 2.90) which is greater than the P content of most organic fertilizers used by small farmers in Zimbabwe. Russian Comfrey liquid manure had the highest K content (3.90%) than all the liquid manures.

All three Liquid Manures had equal Zinc contents. Besides Sea Weed extract few studies have documented the trace element content of Liquid Manures.

5. Acknowledgements

The authors acknowledge the kind assistance rendered by ZOPPA.

REFERENCES

1. IFOAM, 2008. *Definition of Organic Farming*.
http://www.ifoam.org/growing_organic/definitions/doi/index.html. Retrieved 2011-09-3.
2. Gunnarsson, D. (1998). *Heavy metals in fertilisers. Do they cause environmental and health problems*. Fert. Agri. No.85
3. Saxena, S.K., Kavita, M and Arif, A. (1988). Environmental implication of nitrogenous fertilisers industries in India. *Proceedings National Symposium of Environmental Risk Assessment and Management, Valvada, Gujarat*, in December, 1988
4. Chand, S., Anwar, M. & Patra, D.D. (2006). *Influence of long-term application of organic and inorganic fertilizer to build up soil fertility and nutrient uptake in mint-mustard cropping sequence*.

- Communications in Soil Science and Plant Analysis, Canada
5. FAO, 1999. *Organic Farming*. Committee on Agriculture. 15th Session. Rome, 25-29 January 1999.
<http://www.fao.org/./x0075E.htm>. Retrieved 2011-09-30
6. Certified Organic, 2007. *Liquid Organic Fertilisers: friend or foe*. Spring, 2007. www.organic-center.org. Retrieved 2011-09-1.
7. Card, A., Whiting, D., Wilson, C., Reeder, J. (2009) *Organic Fertilisers*. Colorado Gardener Training Certificate. Colorado State University, 2009.
8. Shukla, K. And Pandey, J. (1988). Effect of organic and inorganic soil amendments on rice in saline and sodic soils. *Indian J. Agron*
9. Ramesh, P., Mohan Singh & Subba Rao., A. (2005). Organic farming: Its relevance to the India context. *Current Science*, **88** (4), 561-568.
10. Bokhtiar, S.M. & Sakurai, K. (2005). *Effects of organic manure and chemical fertilizer on soil fertility and productivity of plant and ratoon crops of sugarcane*. Archives of Agronomy and Soil Science.
11. Bioherb (1994). *Training module for Tropical and Subtropical Organic Farming*. International Organic Agriculture, 1994.
12. Giller, K.E and Mapfumo, P. (2001). Soil Fertility Management Strategies and Practises by Smallholder Farmers in Zimbabwe. *African Crop Science Journal Vol. 9 (23)*, pp. 629-644.
13. Harrison, J. (2010). Values of Compost and manures etc. www.allotment.org.uk. Retrieved 2011-09-1.
14. Divers 1999. *Biodynamic Farming and Compost Preparation*. Alternative farming system series. ATTRA.
15. Thirumaran, G., Arumugam, M., Arumugam, R., Anantharaman, P. Effect of Seaweed Liquid Fertilizer on Growth and Pigment Concentration of *Abelmoschus esculentus* (l) medikus *American-Eurasian Journal of Agronomy* 2 (2): 57-66, 2009
16. Crouch, I.J. and J. Van Staden, 1993. Evidence for the presence of growth regulator in commercial seaweed product. *Plant Growth Regulators*, 13: 21-29.
17. Palekar, S., 2006. *Text book on Shoonya Bandovalada naisargika Krushi*. Published by

- Swamy Anand, Agri Prakashana, Bangalore.
18. Vasanthkumar H.H.A., 2006, *Jeevamrut slurry preparation*, Siri Samruddhi, pp.4-5.
 19. Wistinghausen, E., Sattler, F. (1989) *Der landwirtschaftliche Betrieb*. Verlag Eugen Ulmer, Stuttgart 1989
 20. Divers, S. Notes on Compost teas. <http://www.attra.org/attra-pub/PDF/compost-tea-notes.pdf>
 21. The African Organic Farming Foundation. 2004. *The Need for AOFF Projects*.<http://www.africanorganics.org>. Retrieved 2011-9-3)
 22. Sivotwa, E., Baipai, R., Jiyane, J. Organic farming in the small holder farming sector of Zimbabwe. *EJEAFChe*, 6 (2), 2007. [1820-1827]
 23. Hills, L.D (2003) *.Plant studies, Comfrey Past Present and the Future*, Ignoa Publishers, India
 24. Moyo, N.A.G. (1997). *Lake Chivero polluted Lake*, University of Zimbabwe Publications
 25. Satyakala. G, Begum. S, Hasan.A, Jamil.K (1986).Effect of aquatic pollution on the mineral accumulation in water hyacinth. *Proceedings: Plant Sciences. Volume 96, Number 1 / March, 1986*
 26. Lainini.T.W. Redroot Pigweed (*Amaranthus retroflexus*).*Weed Identification*.5. Pennsylvania state University, 1999.
 27. Organics, 2007. Homemade liquid fertilizers. Technical Series. Oregon State University.
 28. Hills,L.D (2003) *.Plant studies, Comfrey Past Present and the Future*,Ignoa Publishers, India
 29. Trinidad, R. T and de Rodrigo López, E.M. 2008.The Water Hyacinth, *Eichhornia crassipes*: an invasive plant in the Guadiana River Basin ,Spain. *Aquatic Invasions (2008) Volume 3, Issue 1: 42-53*
 30. Okoye P.C, Daddy.F and Jlesanmj.B.D. *The nutritive value of water hyacinth (Eichhornia crassipes) and its utilisation in fish feed*. National Institute for Freshwater Fisheries Research,, New Bussa, Niger State.
 31. Akubugwo, I. E., Obasi, N. A., Chinyere G. C. and Ugbogu A. E. Nutritional and chemical value of *Amaranthus hybridus* L. leaves from Afikpo, Nigeria. *African Journal of Biotechnology* Vol. 6 (24), pp. 2833-2839
 32. Humanure Handbook, *Chapter 3: The Carbon / Nitrogen Ratio*. Jenkins Publishing, Grove City, PA. <http://www.jenkinspublishing.com/>
 33. Havlin, J.L and Tisdale, L. Soil fertility and fertilizers: An Introduction to Nutrient Management (7th Edition).
 34. Tukey, H.B. and Wittwer, S.H., 1956. *The entry of nutrients into plants through stem, leaf and fruit, as indicated by radioactive isotopes*. Progress in Nuclear Energy Biological Sciences Series Six, pp.106-114. McGraw-Hill, NY, and Permagon Press, Londo

Experimental Investigation on the Role of MQL and its Effects on Tool Wear in Turning of mild Steel

N.B.BORKAR¹, Dr.S.S.CHAUDHARI², A.B.AMALE³

*(Department of Mechanical Engg. Yashwantrao Chavan College of Engineering Nagpur, Maharashtra., India).

**(Department of Mechanical Engg. Yashwantrao Chavan College of Engineering Nagpur, Maharashtra. India.)

***(Department of Mechanical Engg. Yashwantrao Chavan College of Engineering Nagpur, Maharashtra ,India.)

ABSTRACT

In present study it has been surveyed that, the metal industries using the cutting fluid has become more problematic in terms of both employee health and environmental pollution. But the use of cutting fluid generally causes economy of tools and it becomes easier to keep tight tolerances and to maintain work piece surface properties without damages. Because of them some alternatives has been sought to minimize or even avoid the use of cutting fluid in machining operations. The preferable alternative selected for the present work is the machining with minimum quantity of lubrication (MQL). The growing demands for high productivity of machining need use of high cutting velocity and feed rate. Such machining inherently produces high cutting temperature, which not only reduces tool life but also impairs the product quality. Metal cutting fluids changes the performance of machining operations because of their lubrication, cooling, and chip flushing functions. Hence some alternatives have been sought to minimize or even avoid the use of cutting fluid in machining operations. Some of these alternatives are dry machining and machining with minimum quantity of lubrication (MQL). There was no formation of built up edge at tool due to reduction in the heat and flushing of the chips away from the tool edge. The MQL can able to subsidize the heat generated. This paper deals with experimental investigation on the role of MQL on cutting temperature, tool wear in turning of mild steel

at industrial speed-feed combinations by H.S.S cutting tool. The encouraging results include significant reduction in cutting temperature, tool wears by MQL mainly through favorable chip-tool and work-tool interaction.

Keywords – Chip-tool ,Dry machining , H.S.S, MQL, Work-tool interface etc.

1.Introduction

High production machining of steel inherently generates high cutting zone temperature. Such high temperature causes dimensional deviation and serious failure of cutting tools. It also impairs the surface integrity of the product by inducing tensile residual stresses and surface and subsurface micro cracks in addition to rapid oxidation and corrosion [1,3,4]. Currently, this problem is tried to be controlled by reducing heat generation and removing heat from the cutting zone through optimum selection of machining parameters, proper cutting fluid application and using heat resistant cutting tools. For the companies, the costs related to cutting fluids represent a large amount of the total machining costs. Several research workers state that the costs related to cutting fluids are frequently higher than those related to cutting tools. Consequently, elimination on the use of cutting fluids, if possible, can be a significant economic incentive .In machining, conventional cutting fluid application fails to penetrate the chip- tool interface and thus cannot remove heat effectively. The high pressure jet of soluble oil, when applied at the chip-tool interface, could reduce cutting temperature and improve tool life to certain extent [2,3]. MQL is assisting on the economical front. Dry machining operations are now of great interest and, actually, they meet with success in the field of environmentally friendly manufacturing [6]. In reality, however, they are sometimes less effective when higher machining efficiency, better surface finish quality and severe cutting conditions are required. For these situations, semi-dry operations utilizing very small amounts of cutting lubricants are expected to become a powerful tool and, in fact, they already play a significant role in a number of

practical applications [7,8]. Minimum quantity lubrication refers to the use of cutting fluids of only a minute amount-typically of a flow rate of 50 to 180 ml/hour which is about three to four orders of magnitude lower than the amount commonly used in flood cooling condition, where, for example, up to 10 liters of fluid can be dispensed per minute. The concept of minimum quantity lubrication, sometimes referred to as near dry lubrication [9] or micro-lubrication [10], has been suggested since a decade ago as a means of addressing the issues of environmental intrusiveness and occupational hazards associated with the airborne cutting fluid particles on factory shop floors. The minimization of cutting fluid also leads to economical benefits by way of saving lubricant costs and work -piece/tool/machine cleaning cycle time. The present work experimentally investigates the role of MQL on surface roughness, tool flank wear in turning at different speed combinations by high speed steel tool.

2 .Minimum Quantity Lubrication

2.1 Definition and Concept

The MQL needs to be supplied at high pressure and impinged at high speed through the spray painting gun on the cutting zone. Considering the conditions required for the present work and uninterrupted supply of MQL at a constant pressure of around 6 bar over a reasonably long cut, a MQL delivery system was designed, fabricated and used. The schematic view of the MQL set-up is shown in Fig 1.

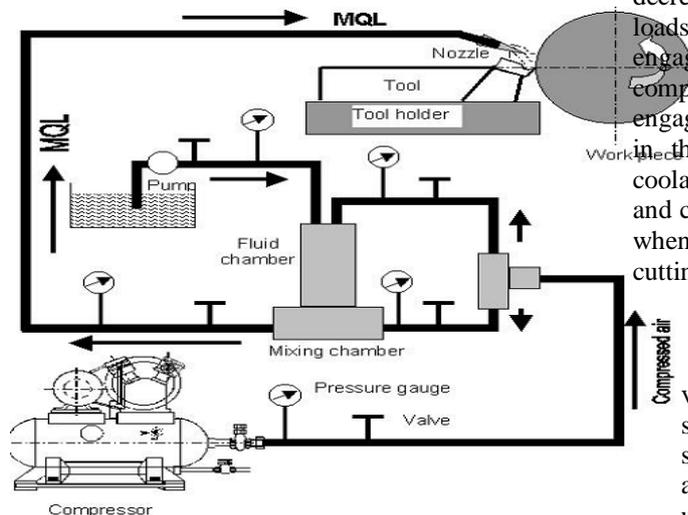


Figure .1Working of

MQL system

cutting tools. Consequently, elimination on the use of cutting fluids, if possible, can be a significant economic incentive. Considering the high cost associated with the use of cutting fluids and projected escalating costs when the stricter environmental laws are enforced, the choice seems obvious. Because of them some alternatives has been sought to minimize or even avoid the use of cutting fluid in machining operations. Some of these alternatives are dry machining and machining with minimum quantity lubrication (MQL).

2.1.2 Working of MQL

High cutting zone temperature is generally tried to be controlled by employing flood cooling by soluble oil. In high speed-feed machining, conventional cutting fluid application fails to penetrate the chip-tool interface and thus cannot remove heat effectively and the use of cutting fluid has become more problematic in terms of both employee health and environmental pollution. Addition of extreme pressure additives in the cutting fluids does not ensure penetration of coolant at the chip-tool interface to provide lubrication and cooling . Minimum quantity lubrication (MQL) is based on the principle that a drop of liquid is split by an air flow, distributed in streaks and transported in the direction of flow of air. The consumptions of oil in industrial applications are in the range of the length of machining was increased wear by the tool increased. MQL consists of a mixture of pressurized air and oil micro- droplets applied directly into the interface between the tool and chips. However, the question of how the lubricants can decrease the friction under very high temperature and loads is still not answered especially for long engagements times. MQL decreased the contact length compared to dry cutting for both short and long engagement time. Addition of extreme pressure additives in the cutting fluids does not ensure penetration of coolant at the chip-tool interface to provide lubrication and cooling . However, high-pressure jet of vegetable oil, when applied at the chip-tool interface, could reduce cutting temperature and improve tool life to some extent.

In MQL machining, a small amount of vegetable oil or biodegradable synthetic ester is sprayed to the tool tip with compressed air. The mild steel work-piece machining with MQL arrangement as shown in fig.2, and fig.3 as the experimental set up.

Several research workers state that the costs related to cutting fluids are frequently higher than those related to



Figure.2 Photographic view of the experimental set-up for turning steel with MQL operation

Figure.3 Photographic view of mixing cup for turning operation

MQL machining is nearly equal or often better than the traditional wet machining in tool life and surface finish when cutting steels and aluminum alloys. A recent survey conducted on the production of the European automotive industry revealed that the expense of cooling lubricant comprises nearly 20% of the total manufacturing cost is referred in fig .4.

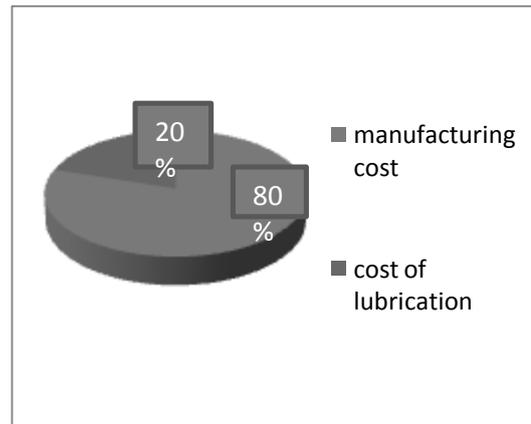


Figure.4 percentage disk for lubrication cost

In comparison to cutting tools, the cooling lubricant cost is significantly higher. As a result, the need to reduce cutting fluids consumption is strong. Furthermore, the permissible exposure level for metal-working fluid aerosol concentration is 5 mg/m³, according to the U.S. Occupational Safety and Health Administration (OSHA) [17], and is 0.5 mg/m³ according to the U.S. National Institute for Occupational Safety and Health (NIOSH) [17]. The oil mist level in U.S. automotive parts manufacturing facilities has been estimated to be generally on the order of 20-90 mg/m³ with the use of traditional flood cooling and lubrication [10]. This suggests an opportunity for improvement of several orders of magnitude. The minimal quantity lubrication (MQL) system is probably the most representative application of semi-dry machining. The purpose of this research work is to experimentally investigate the influence of MQL on cutting tool wear in turning mild steel at industrial speed-feed conditions by H.S.S. tool and compare the effectiveness of MQL with that of dry and wet machining.

3.EXPERIMENTAL CONDITIONS AND PROCEDURE

The present investigation used, mild steel of initial diameter (size: Ø20mm×80 mm) is plain turning, 3hp Lathe HSS (Miranda S-500) at industrial speed-feed combinations under dry, wet and MQL conditions. The experimental conditions are given in Table-1. The ranges of the cutting velocity (V_c) and feed rate (S_o) were selected based on the tool manufacturer's recommendation and industrial practices. Depth of cut, being a less significant parameter, was kept fixed.

Table1. : Experimental conditions

| | |
|------------------------|---|
| Machine tool: | 3hp Lathe |
| Work-piece: | Mild Steel (size: Ø20mm× mm) |
| Cutting tool | HSS ,Miranda S-500 |
| Working tool ometry | 8°, 8°, 6°, 6°, 8°, 12°, 0.6 m) |
| Cutting velocity, Vc | 63, 80, 93 m/min |
| Feed rate, So: | 0.035,0.050 and 0.820 mm/rev |
| Depth of cut, t: | 0.8,0.85 and 1.0 mm |
| MQL supply: | Air: 6 bar, Lubricant: 200 ml/h |
| Environment: | Wet (flood cooling) Minimum Quantity lubrication (MQL) |

3.1 Tool Wear

Productivity and economy of manufacturing by machining are significantly influenced by life of the cutting tools. Cutting tools may fail by brittle fracturing, plastic deformation or gradual wear. Turning steel having enough strength, toughness and hot hardness generally fail by gradual wears. With the progress of machining the tools attain crater wear at the rake surface and flank wear at the clearance surfaces, as schematically shown in Fig.3 due to continuous interaction and rubbing with the chips and the work surfaces, respectively.

3.1.1 Crater Wear

The crater wear is usually avoided by selecting a cutting speed and tool material that does not have an affinity to diffusion with the work material. The flank wear, on the other hand, leads to loss of cutting edge, and affects the dimension and surface finish quality. Crater wear occurs on the rake face of the tool where the chip moves under friction and normal loads at elevated temperatures, leading to wear. Since all cutting edges have a finite sharpness, the friction between the flank face of the cutting tool and the freshly cut work surface causes flank wear. The cutting tool wear is well-known affecting the tool life and the surface quality of the finished product. When tool wear is beyond a certain threshold, the tool fails catastrophically due to excessive

stresses and rising thermal within the tool edge caused by large friction forces.

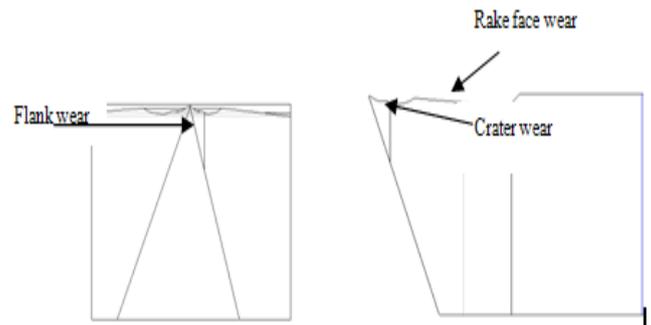


Figure.5 Photographic view of the crater wear

the MQL has reduced tool wear. remarkably in machining mild steel. This is reasonably attributed to extremely lubrication provided by the MQL jet, which could, at least partially, reach the work-tool interfaces, unlike chip-tool interface is shown in fig.4. The deep grooves parallel to the cutting edges of the insert are likely to help entry of larger fraction of the MQL jet at the flank surfaces.

3.1.2 Flank wear

Among the aforesaid wears, the principal flank wear is the most important because it raises the cutting forces and the related problems. The life of carbide tools, which mostly fail by wearing, is assessed by the actual machining time after which the average value (VB) of its principal flank wear reaches a limiting value, like 0.3 mm. Therefore, attempts should be made to reduce the rate of growth of flank wear (VB) in all possible ways without sacrifice in MRR. During the machining process, the cutting tools are loaded with the heavy forces resulting from the deformation process in chip formation and friction between the tool and work-piece. The heat generated at the deformation and friction zones overheats the tool, the chip and partially the work-piece. All the contact surfaces are usually clean and chemically very active; therefore the cutting process is connected with complex physical-chemical processes. Wear on the tool, which occurs as the consequence of such processes, is reflected as progressive wearing of particles from the tool surface.

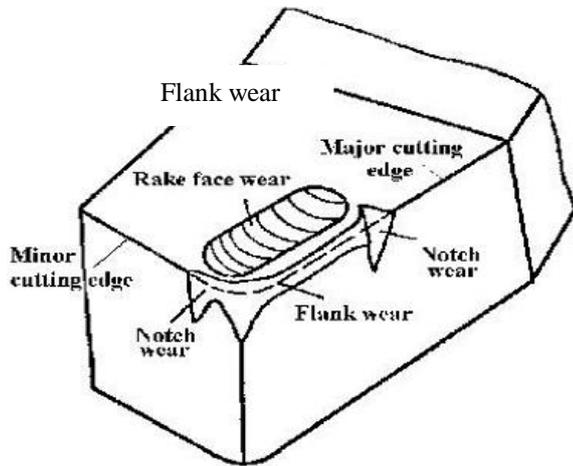


Figure.6 Photographic view of the

flank wear

In machining research, a cutting tool is generally said to have failed when its V reaches a specific value, mostly 0.3 mm. It is very important to note in Fig.4 that tool life has improved from 31 min to 48 min, i.e. almost by one and half times increase in tool life have been possible by MQL. Fig.4 also depicts how flank notch wear, V remarkably decreased due to MQL. Deep notching, if forms, not only raises cutting forces but also may cause catastrophic tool failure prematurely and randomly, which is extremely harmful and undesirable for the present days' sophisticated and expensive manufacturing systems. So, proper MQL is expected also to enhance reliability and safety of machining processes and systems.

4.Result and Discussion

MQL in the form of thin but high speed was impinged from a specially designed nozzle along the auxiliary cutting edge of the H.S.S. cutting tool and mild steel round work piece interface . The MQL jet has been used mainly to target the rake surface and flank surface along the auxiliary cutting edge and to protect the auxiliary flank to enable better dimensional accuracy. MQL is expected to provide some favorable effects mainly through reduction in cutting temperature by making the thermal conductive film at work-tool chip interface at experimental trial on wet and MQL condition . With the S/N and ANOVA analyses, the optimal combination of the process parameters can be predicted. Finally, a confirmation experiment is conducted to verify the

optimal process parameters obtained from the parameter design. In this paper, the cutting parameter design by the Taguchi method is adopted to obtain optimal machining performance in turning. The reliable tool-work minimum quantity lubrication technique with proper combination has been employed to measure the average tool wear, surface roughness during turning of mild steel by the HSS as shown in table.1,table.2,table.3.

4.1.Experimentation results of mild steel turning tool wear dry ,wet , MQL condition.

Table 1. Experimentation results of Dry condition

| EX. No. | Cutting Speed (m /min) | Feed rate (mm /rev) | Dept of cut (mm) | Tool wear V _B ,mm |
|---------|------------------------|---------------------|------------------|------------------------------|
| 1 | 63 | 0.16 | 0.8 | 0.2 |
| 2 | 80 | 0.55 | 0.85 | 0.24 |
| 3 | 93 | 0.24 | 1.0 | 0.31 |

Table 2. Experimentation results of Wet condition

| EX. No. | Cutting Speed (m /min) | Feed rate (mm /rev) | Dept of cut (mm) | Tool wear V _B ,mm |
|---------|------------------------|---------------------|------------------|------------------------------|
| | 63 | 0.16 | 0.8 | 0.35 |
| 2 | 80 | 0.55 | 0.85 | 0.39 |
| 3 | 93 | 0.24 | 1.0 | 0.43 |

Table 3. Experimentation results of MQL condition

| EX. No. | Cutting Speed (m /min) | Feed rate (mm /rev) | Dept of cut (mm) | Tool wear V _B ,mm |
|---------|------------------------|---------------------|------------------|------------------------------|
| 1 | 63 | 0.16 | 0.8 | 0.21 |
| 2 | 80 | 0.55 | 0.85 | 0.19 |
| 3 | 93 | 0.24 | 1.0 | 0.24 |

4.2 Graphical representation of an experimental results for dry , wet , MQL machining

The reliable tool-work minimum quantity lubrication technique with proper combination has been employed to measure the average tool wear, surface roughness during turning of mild steel by the HSS as shown in figure.7and figure8.

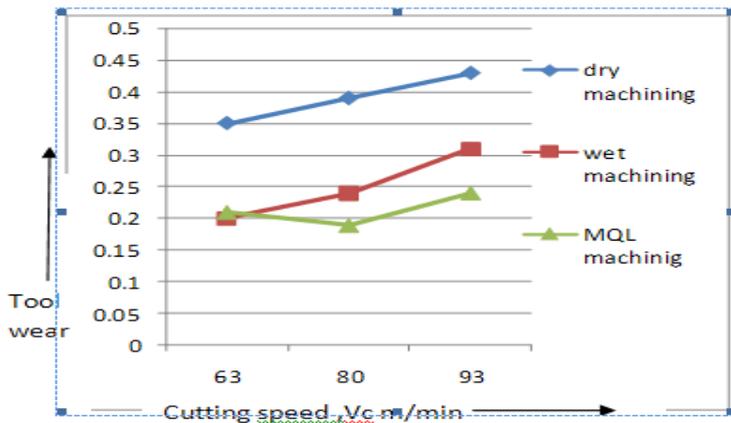


Figure. 7Graphical representation of cutting speed Vs tool wear.

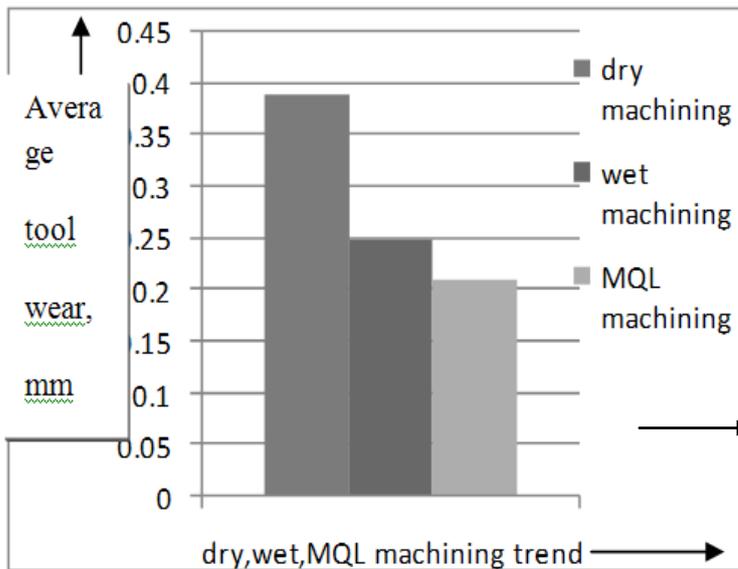


Figure.8 graphical relationship between dry ,wet ,MQL machining.

4. CONCLUSIONS

The present MQL systems enabled reduction in average tool wear for reduction and significant improvement in the major machinability. The most significant contribution of application of MQL in machining the steel by H.S.S tool undertaken has been the high reduction in flank wears, which would enable remarkable improvement in tool life. Such reduction in tool wear might have been possible for retardation of abrasion and notching, decrease or prevention of adhesion and diffusion type thermal sensitive wear at the flanks and reduction of built-up edge (BUE) formation which accelerates wear at the cutting edges by chipping and flaking. Deep notching and grooving, which are very detrimental and may cause premature and catastrophic failure of the cutting tools, are remarkably reduced by MQL.MQL provided better surface finish ultimately increases dimensional accuracy, substantially improved mainly due to significant reduction of wear and damage at the tool tip by the application of MQL. The cutting performance of MQL machining is better than that of conventional machining with flood cutting fluid supply.

References

- [1] N. R. Dhar, M.W. Islam “ A Study of Effects of MQL on Tool Wear, Job Dimension and Finish in Turning MILD STEEL Steel” AESEAP Journal of Engineering Education 2007; Vol. 31, No. 2
- [2] S. Thamizhmanii *, Rosli, S.Hasan”study of mql on inconel 718 steel “
- [3] Faculty of Mechanical and Manufacturing Engineering, University Tun Hussein Onn Malaysia 86400, Parit Raja, Batu Pahat, Johor, Malaysia, Received 28.01.2009; published in revised form 01.09.2009
- [4] P. Leskover and J.Grum, “The Metallurgical Aspect of Machining”, Ann. CIRP 35, pp. 537-550, 1986
- [5] H.K. Tonshoff and E. Brinkomeier, “Determination of the Mechanical and Thermal Influences on Machined Surface by Microhardness and Residual Stress Analysis”, Ann. CIRP 29, pp.519-532, 1986
- [6] C. Cassin and G. Boothroyd, “Lubrication Action of Cutting Fluids”, J. Mech.Eng. Science, 7(1) pp. 67-81, 1965
- [7] F. Klocke and G. Eisenblatter, “Dry Cutting”, Annals of the CIRP, 46/2, pp.519-526, 1997
- [8] U. Heisel, et al., “Application of Minimum Quantity Cooling Lubrication Technology in Cutting Processes”, Prod. Eng, Vol. II/1, pp.49-54, 1994
- [9] J. W. Sutherland, et al., “An Experimental Investigation of Air Quality in Wet and Dry Turning”, Ann. of CIRP, 49/1, pp.61-64, 2000
- [9] J. McCabe and M. A. Ostaraff, “Performance

Experience with Near-Dry Machining of Aluminum", *Lubr. Eng.*, 57/12, pp. 22-27, 2001

[10]. S. Suda, et al., "Evaluation of Machinability with MQL System and Effectiveness in Production Lines", *Proceedings of the International Tribology Conference, Nagasaki*, pp.203-208, 2001

[11]T. F. MaClure, R. Adams and M. D. Gugger, "Comparison of Flood vs. Micro-lubrication on Machining Performance, website: <http://www.unist.com/techsolve.html> , 2001.

[12]N. Richard ,D. Aspiwall use of ceramic tool for machining nickel base alloys *international journal of engineering and tool manufacturing* 296 (1989).

[13]T. Brockhoff and A. Walter, "Fluid Minimization in Cutting and Grinding, Abrasives", pp. 38-42, 1998

[14]R. B. Aronson, "Why Dry Machining", *Manufacturing Engineering*, pp. 33-36, 1995

[15]U.S. Department of Health and Human Services, "Occupational Exposure to Metalworking Fluids", NIOSH Publication No. 98-102, January 1998

[16]E. O. Bennett and D. L. Bennett, "Occupational Airway Diseases in the Metalworking Industry", *Tribology International*, 18/3, pp. 169-176, 1998

Potential of a Low Heat Rejection Diesel Engine with Crude Pongamia Oil

Chennakesava Reddy¹, M.V.S. Murali Krishna^{2*}, P.V.K.Murthy³ and T. Ratna Reddy⁴

¹Department of Mechatronics, Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad – 500 075

^{2,4} Mechanical Engineering Department, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad-500 075, Andhra Pradesh, India, *E-mail: maddalivs@gmail.com

³ Vivekananda Institute of Science and Information Technology, Shadnagar, Mahabubnagar-509216, Andhra Pradesh, India,

ABSTRACT

Investigations are carried out to evaluate the performance of a low heat rejection (LHR) diesel engine consisting of air gap insulated piston with 3-mm air gap, with superni 2/4 (an alloy of nickel) crown and air gap insulated liner with superni insert with different operating conditions of crude pongamia oil (CPO) with varied injection pressure and injection timing. Performance parameters are determined at various magnitudes of brake mean effective pressure. Pollution levels of smoke and oxides of nitrogen (NO_x) are recorded at the peak load operation of the engine. Combustion characteristics at peak load operation of the engine are measured with TDC (top dead centre) encoder, pressure transducer, console and special pressure-crank angle software package. Conventional engine (CE) showed deteriorated performance, while LHR engine showed improved performance with CPO operation at recommended injection timing and pressure and the performance of both version of the engine is improved with advanced injection timing and at higher injection pressure when compared with CE with pure diesel operation. Peak brake thermal efficiency increased by 5%, smoke levels decreased by 4% and NO_x levels increased by 40% with CPO operation on LHR engine at its optimum injection timing, when compared with pure diesel operation on CE at manufacturer's recommended injection timing.

Keywords: Crude pongamia oil, LHR engine, Performance, Pollution levels, Combustion characteristics.

1. INTRODUCTION

In the scenario of increase of vehicle population at an alarming rate due to advancement of civilization, use of diesel fuel is not only transport

sector but also in agriculture sector leading to fast depletion of diesel fuels and increase of pollution levels with these fuels, the search for alternate fuels on has become pertinent for the engine manufacturers, users and researchers involved in the combustion research. It is well known fact that about 30% of the energy supplied is lost through the coolant and the 30% is wasted through friction and other losses, thus leaving only 30% of energy utilization for useful purposes. In view of the above, the major thrust in engine research during the last one or two decades has been on development of LHR engines. Several methods adopted for achieving LHR to the coolant are i) using ceramic coatings on piston, liner and cylinder head ii) creating air gap in the piston and other components with low-thermal conductivity materials like superni, cast iron and mild steel etc. Through ceramic coatings provided insulation and improved brake specific fuel consumption (BSFC), peeling of coating was reported by various researchers [1-4] after certain hours of duration. Creating an air gap in the piston involved the complications of joining two different metals. Though Parker et al. [5] observed effective insulation provided by an air gap, the bolted design employed by them could not provide complete sealing of air in the air gap. Dhinagar et al. [6] applied different degrees of insulation like ceramic coated cylinder head, air gap insulated piston and air gap insulated liner and conducted experiments with pure diesel operation and reported LHR version of the engine improved the performance. Rama Mohan [7] made a successful attempt of screwing the crown made of low thermal conductivity material, nimonic (an alloy of nickel) to the body of the piston, by keeping a gasket, made of nimonic, in between these two parts. Low degree of insulation provided by these researchers

[7] was not able to burn low cetane fuels of vegetable oils.

The idea of using vegetable oil as fuel has been around from the birth of diesel engine. Rudolph diesel, the inventor of the engine that bears his name, experimented with fuels ranging from powdered coal to peanut oil. Experiments were conducted [11-15] with CE with vegetable oils and reported that performance was deteriorated with CE.

Bhaskar et al. [17] conducted experimental investigations on jatropha oil with LHR engine, which consisted of ceramic-coated cylinder head and air gap cylinder liner and reported that performance was improved and pollution levels of hydrocarbon and smoke decreased with LHR version of the engine with jatropha oil when compared with CE with pure diesel operation. Ignition improvers to jatropha oil further improved the performance and reduced the pollution levels. S.Jabez Dhinagar et al [18] tested three vegetable oils, neem oil, rice bran oil and karanja oil in LHR engine and reported that performance of vegetable oils was improved with pre-heating. Hanbey Hazar [19] conducted investigations on LHR engine with the cylinder head, exhaust, and inlet valves of a diesel engine were coated with the ceramic material MgO-ZrO₂ by the plasma spray method, while the piston surface was coated with ZrO₂. Thus, a thermal barrier was provided for the elements of the combustion chamber with these coatings. Using identical coated and uncoated engines, the effects of canola methyl ester produced by the transesterification method, and ASTM No. 2D fuel on engine performance and exhaust emissions were studied. Tests were performed on the uncoated engine, and then repeated on the coated engine and the results were compared. An increase in engine power and decrease in specific fuel consumption, as well as significant improvements in exhaust gas

emissions and smoke density, were observed for all test fuels used in the coated engine compared with that of the uncoated engine. Rajendra Prasath et al. [20] conducted experiments on LHR engine with partially stabilized zirconia coating on inside portion of cylinder head with bio-diesel and reported that performance was improved with LHR engine.

Little literature is available in evaluating the performance of LHR engine with air gap piston and air gap liner with superni (an alloy of nickel, a low thermal conductivity material) with varying engine parameters at different operating conditions of the vegetable oil. The present paper attempts to evaluate the performance of LHR engine, which contains air gap piston with superni crown and air gap insulated liner with superni insert with different operating conditions of CPO with varying engine parameters of change of injection pressure and timing and compared with CE at recommended injection timing and injection pressure.

2.EXPERIMENTAL PROGRAMME

Fig.1 gives the details of insulated piston and insulated liner employed in the experimentation. LHR diesel engine contains a two-part piston; the top crown made of low thermal conductivity material, superni-90 screwed to aluminum body of the piston, providing a 3mm-air gap in between the crown and the body of the piston. The optimum thickness of air gap in the air gap piston is found to be 3-mm [7], for better performance of the engine with superni inserts. A superni-90 insert is screwed to the top portion of the liner in such a manner that an air gap of 3-mm is maintained between the insert and the liner body. At 500°C the thermal conductivity of superni-90 and air are 20.92 and 0.057 W/m-K respectively. The properties of vegetable oil are shown in Table-1.

TABLE-1.
PROPERTIES OF THE NON-EDIBLE VEGETABLE OIL AND DIESEL

| Test Fuel | Viscosity at 25 ° C (centi-poise) | Density at 25 ° C | Cetane number | Calorific value (kJ/kg) |
|----------------------|---|----------------------|------------------|-------------------------------|
| Diesel | 12.5 | 0.84 | 55 | 42000 |
| Pongamia oil (crude) | 125 | 0.91 | 48 | 37100 |

Experimental setup used for the investigations of LHR diesel engine with crude pongamia oil (CPO) is shown in Fig.2. CE has an aluminum alloy

piston with a bore of 80 mm and a stroke of 110mm. The rated output of the engine is 3.68 kW at a rate speed of 1500 rpm. The compression ratio

is 16:1 and manufacturer's recommended injection timing and injection pressures are 27°bTDC and 190 bar respectively. The fuel injector has 3-holes of size 0.25-mm. The combustion chamber consists of a direct injection type with no special arrangement for swirling motion of air. The engine is connected to electric dynamometer for measuring brake power of the engine. Burette method is used for finding fuel consumption of the engine. Air-consumption of the engine is measured by air-box method. The naturally aspirated engine is provided with water-cooling system in which inlet temperature of water is maintained at 60°C by adjusting the water flow rate. Engine oil is provided with a pressure feed system. No temperature control is incorporated, for measuring the lube oil temperature. Copper shims of suitable size are provided in between the pump body and the engine frame, to vary the injection timing and its effect on the performance of the engine is studied, along with the change of injection pressures from 190 bar to 270 bar (in steps of 40 bar) using nozzle testing device. The maximum injection pressure is restricted to 270 bar due to

practical difficulties involved. Exhaust gas temperature (EGT) is measured with thermocouples made of iron and iron-constantan. Pollution levels of smoke and NO_x are recorded by AVL smoke meter and Netel Chromatograph NO_x analyzer respectively at the peak load operation of the engine. Piezo electric transducer, fitted on the cylinder head to measure pressure in the combustion chamber is connected to a console, which in turn is connected to Pentium personal computer. TDC encoder provided at the extended shaft of the dynamometer is connected to the console to measure the crank angle of the engine. A special P-θ software package evaluates the combustion characteristics such as peak pressure (PP), time of occurrence of peak pressure (TOPP), maximum rate of pressure rise (MRPR) and time of occurrence of maximum rate of pressure rise (TOMRPR) from the signals of pressure and crank angle at the peak load operation of the engine. Pressure-crank angle diagram is obtained on the screen of the personal computer

3.RESULTS AND DISCUSSION

3.1 Performance Parameters

The variation of brake thermal efficiency (BTE) with brake mean effective pressure (BMEP) in the conventional engine (CE) with CPO, at various injection timings at an injection pressure of 190 bar, is shown in Fig.3. The variation of BTE with BMEP with pure diesel operation on CE at recommended injection timing is also shown for comparison purpose. CE with vegetable oil showed the deterioration in the performance for entire load range when compared with the pure diesel operation on CE at recommended injection timing. Although carbon accumulations on the nozzle tip might play a partial role for the general trends observed, the difference of viscosity between the diesel and vegetable oil provided a possible explanation for the deterioration in the performance of the engine with vegetable oil operation. The amount of air entrained by the fuel spray is reduced, since the fuel spray plume angle is reduced, resulting in slower fuel- air mixing. In addition, less air entrainment by the fuel spray suggested that the fuel spray penetration might increase and resulted in more fuel reaching the combustion chamber walls. Furthermore droplet mean diameters (expressed as Sauter mean) are larger for vegetable oil leading to reduce the rate of heat release as compared with diesel fuel. This also, contributed the higher ignition (chemical) delay of the vegetable oil due to lower cetane number. According to the qualitative image of the combustion under the crude vegetable oil operation

with CE, the lower BTE is attributed to the relatively retarded and lower heat release rates. BTE increased with the advancing of the injection timing in CE with the vegetable oil at all loads, when compared with CE at the recommended injection timing and pressure. This is due to initiation of combustion at earlier period and efficient combustion with increase of air entrainment in fuel spray giving higher BTE. BTE increased at all loads when the injection timing is advanced to 32°bTDC in the CE at the normal temperature of vegetable oil. The increase of BTE at optimum injection timing over the recommended injection timing with vegetable oil with CE could be attributed to its longer ignition delay and combustion duration. BTE increased at all loads when the injection timing is advanced to 32°bTDC in CE, at the preheated temperature of CPO. That, too, the performance is improved further in CE with the preheated vegetable oil for entire load range when compared with normal vegetable oil. Preheating of the vegetable oil reduced the viscosity, which improved the spray characteristics of the oil and reduced the impingement of the fuel spray on combustion chamber walls, causing efficient combustion thus improving BTE.

The variation of BTE with BMEP in the LHR engine with CPO, at various injection timings at an injection pressure of 190 bar, is shown in Fig.4. LHR version of the engine showed the marginal improvement in the performance for entire load

range compared with CE with pure diesel operation. High cylinder temperatures helped in better evaporation and faster combustion of the fuel injected into the combustion chamber. Reduction of ignition delay of the vegetable oil in the hot environment of the LHR engine improved heat release rates and efficient energy utilization. Preheating of vegetable oil improves performance further in LHR version of the engine. The optimum injection timing is found to be 30°bTDC with LHR engine with normal CPO. Since the hot combustion chamber of LHR engine reduced ignition delay and combustion duration and hence the optimum injection timing is obtained earlier with LHR

engine when compared with CE with the vegetable oil operation.

Injection pressure is varied from 190 bars to 270 bars to improve the spray characteristics and atomization of the vegetable oils and injection timing is advanced from 27 to 34°bTDC for CE and LHR engine. Table-2 shows the variation of BTE with injection pressure and injection timing at different operating conditions of CPO with different configurations of the engine. BTE increases with increase in injection pressure in both versions of the engine at different operating conditions of the vegetable oil.

TABLE -2
THE VARIATION OF PEAK BTE WITH INJECTION TIMING AND INJECTION PRESSURE IN THE CONVENTIONAL AND LHR ENGINES AT DIFFERENT OPERATING CONDITIONS OF THE VEGETABLE OIL

| Injection Timing (° bTDC) | Test Fuel | Peak BTE (%) | | | | | | | | | | | |
|---------------------------|-----------|---------------------------|-----|------|------|------|------|---------------------------|------|------|------|------|------|
| | | Conventional Engine | | | | | | LHR Engine | | | | | |
| | | Injection Pressure (Bars) | | | | | | Injection Pressure (Bars) | | | | | |
| | | 190 | | 230 | | 270 | | 190 | | 230 | | 270 | |
| | | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT |
| 27 | DF | 28 | -- | 29 | --- | 30 | -- | 29 | -- | 30 | -- | 30.5 | -- |
| | CPO | 25 | 26 | 26 | 27 | 27 | 28 | 29 | 29.5 | 29.5 | 30 | 30 | 30.5 |
| 30 | DF | 29 | --- | 30 | -- | 30.5 | -- | 29.5 | -- | 30.5 | -- | 31 | -- |
| | CPO | 26.5 | 27 | 27 | 27.5 | 28.5 | 29 | 29.5 | 30 | 30 | 30.5 | 30.5 | 31 |
| 31 | DF | 29.5 | -- | 30 | -- | 31 | -- | 30 | -- | 31 | -- | 31 | -- |
| | CPO | 27.5 | 28 | 28.5 | 29 | 28 | 28.5 | -- | -- | -- | --- | -- | -- |
| 32 | DF | 30 | | 30.5 | | 30.5 | | | | | | | |
| | CPO | 28.5 | 29 | 28 | 28.5 | 27.5 | 28 | -- | -- | -- | -- | -- | -- |
| 33 | DF | 31 | | 31 | | 30 | --- | -- | -- | -- | -- | -- | - |

DF-Diesel Fuel, CPO-Crude Pongamia Oil, NT- Normal or Room Temperature , PT- Preheat Temperature

The improvement in BTE at higher injection pressure is due to improved fuel spray characteristics. However, the optimum injection timing is not varied even at higher injection pressure with LHR engine, unlike the CE. Hence it is concluded that the optimum injection timing is 32°bTDC at 190 bar, 31°bTDC at 230 bar and 30°bTDC at 270 bar for CE. The optimum injection timing for LHR engine is 30°bTDC irrespective of injection pressure. Peak BTE is higher in LHR engine when compared with CE with different operating conditions of the vegetable oils.

Fig.5 shows the variation of the exhaust gas temperature (EGT) with BMEP in CE and LHR engine with CPO at normal temperature at the recommended and optimized injection timings at

an injection pressure of 190 bar. CE with CPO at the recommended injection timing recorded higher EGT at all loads compared with CE with pure diesel operation. Lower heat release rates and retarded heat release associated with high specific energy consumption caused increase in EGT in CE. Ignition delay in the CE with different operating conditions of vegetable oil increased the duration of the burning phase. LHR engine recorded lower value of EGT when compared with CE with vegetable oil operation. This is due to reduction of ignition delay in the hot environment with the provision of the insulation in the LHR engine, which caused the gases expand in the cylinder giving higher work output and lower heat rejection. This showed that the performance is improved with

LHR engine over CE with vegetable oil operation. The magnitude of EGT at peak load decreased with advancing of injection timing and with increase of injection pressure in both versions of the engine with vegetable oil. Preheating of the vegetable oil further reduced the magnitude of EGT, compared with normal vegetable oil in both versions of the engine. Table-3 shows the variation of EGT with

injection pressure and injection timing at different operating conditions of CPO with different configurations of the engine. EGT decreases with increase in injection pressure and injection timing with both versions of the engine, which confirms that performance is increased with injection pressure. Preheating of vegetable oil decreases EGT in both versions of the engine.

TABLE-3
THE VARIATION OF EGT AT THE PEAK LOAD WITH INJECTION TIMING AND INJECTION PRESSURE IN THE CONVENTIONAL AND LHR ENGINES AT DIFFERENT OPERATING CONDITIONS OF THE VEGETABLE OIL

| Injection timing (° b TDC) | Test Fuel | EGT at the peak load (°C) | | | | | | | | | | | |
|----------------------------|-----------|---------------------------|-----|-----|-----|-----|-----|---------------------------|-----|-----|-------|------|-----|
| | | Conventional Engine | | | | | | LHR Engine | | | | | |
| | | Injection Pressure (Bars) | | | | | | Injection Pressure (Bars) | | | | | |
| | | 190 | | 230 | | 270 | | 190 | | 230 | | 270 | |
| | | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT |
| 27 | DF | 425 | -- | 410 | --- | 395 | -- | 475 | --- | 460 | -- | 445 | -- |
| | CPO | 525 | 500 | 500 | 490 | 490 | 465 | 470 | 465 | 465 | 460 | 460 | 455 |
| 30 | DF | 410 | --- | 400 | -- | 385 | --- | 455 | --- | 450 | -- | 445 | -- |
| | CPO | 500 | 490 | 490 | 480 | 425 | 415 | 465 | 460 | 460 | 455 | 455 | 450 |
| 31 | DF | 400 | --- | 390 | -- | 375 | --- | 450 | --- | 445 | --- | 440 | --- |
| | CPO | 465 | 460 | 425 | 415 | 435 | 425 | | | | | | |
| 32 | DF | 390 | | 380 | | 380 | | 29 | -- | 30 | -- | 30.5 | -- |
| | CPO | 425 | 415 | 435 | 425 | 445 | 435 | ----- | --- | --- | ---- | --- | - |
| | EPO | 410 | 400 | 400 | 390 | 410 | 400 | -- | - | -- | ----- | --- | -- |
| 33 | DF | 375 | --- | 375 | --- | 400 | -- | -- | -- | -- | --- | -- | -- |
| | EPO | 400 | 390 | 410 | 400 | 420 | 410 | -- | -- | -- | -- | -- | -- |

DF-Diesel Fuel, CPO-Crude Pongamia Oil, NT - Normal or Room Temperature, PT - Preheat Temperature

Fig.6 shows the variation of the volumetric efficiency (VE) with BMEP in CE and LHR engine with CPO at the recommended and optimized injection timings at an injection pressure of 190 bar. VE decreased with an increase of BMEP in both versions of the engine. This is due to increase of gas temperature with the load. At the recommended injection timing, VE in the both versions of the engine with CPO operation decreased at all loads when compared with CE with pure diesel operation. This is due increase of temperature of incoming charge in the hot environment created with the provision of

insulation, causing reduction in the density and hence the quantity of air with LHR engine. VE increased marginally in CE and LHR engine at optimized injection timings when compared with recommended injection timings with CPO. This is due to decrease of un-burnt fuel fraction in the cylinder leading to increase in VE in CE and reduction of gas temperatures with LHR engine. Table-4 shows the variation of VE with injection pressure and injection timing at different operating conditions of CPO with different configurations of the engine.

TABLE-4
THE VARIATION OF VOLUMETRIC EFFICIENCY (VE) AT THE PEAK LOAD WITH THE INJECTION TIMING AND INJECTION PRESSURE IN THE CONVENTIONAL AND LHR ENGINES, AT DIFFERENT OPERATING CONDITIONS OF THE VEGETABLE OIL

| Injection timing (°bTDC) | Test Fuel | Volumetric efficiency (%) | | | | | | | | | | | |
|-----------------------------|-----------|---------------------------|------|------|------|------|------|---------------------------|------|------|------|------|------|
| | | Conventional Engine | | | | | | LHR Engine | | | | | |
| | | Injection Pressure (Bars) | | | | | | Injection Pressure (Bars) | | | | | |
| | | 190 | | 230 | | 270 | | 190 | | 230 | | 270 | |
| | | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT |
| 27 | DF | 85 | -- | 86 | -- | 87 | -- | 78 | -- | 80 | -- | 82 | -- |
| | CPO | 78.5 | 79.5 | 79.5 | 80.5 | 80.5 | 81.5 | 75.5 | 76.5 | 76.5 | 77.5 | 77.5 | 78.5 |
| 30 | DF | 86 | -- | 87 | -- | 88 | --- | 80 | -- | 82 | -- | 83 | -- |
| | CPO | 79 | 80 | 80 | 81 | 81 | 82 | 77 | 77.5 | 78.5 | 79.5 | 79.5 | 80.5 |
| 31 | DF | 87 | -- | 87.5 | -- | 89 | -- | 82 | -- | 83 | -- | 84 | -- |
| | CPO | 79.5 | 80.5 | 80.5 | 81.5 | 81.5 | 82.5 | - | -- | -- | -- | -- | - |
| 32 | DF | 87.5 | -- | 88 | -- | 87 | -- | - | -- | - | -- | -- | - |
| | CPO | 80 | 81 | 81 | 82 | 82 | 83 | -- | -- | -- | -- | -- | -- |
| | EPO | 81 | 82 | 82 | 83 | 83 | 84 | -- | -- | -- | - | --- | -- |
| 33 | DF | 89 | -- | 89 | -- | 86 | -- | -- | -- | -- | -- | -- | -- |

DF-Diesel Fuel, CPO-Crude Pongamia Oil, NT- Normal or Room Temperature , PT- Preheat Temperature

VE increased marginally with the advancing of the injection timing and with the increase of injection pressure in both versions of the engine. This is due to better fuel spray characteristics and evaporation at higher injection pressures leading to marginal increase of VE. This is also due to the reduction of

residual fraction of the fuel, with the increase of injection pressure. Preheating of the vegetable oil marginally improved VE in both versions of the engine, because of reduction of un-burnt fuel concentration with efficient combustion, when compared with the normal temperature of oil.

3.2 POLLUTION LEVELS

Fig.7 shows the variation of the smoke levels with BMEP in CE and LHR engine with vegetable oil operation at the recommended and optimized injection timings at an injection pressure of 190 bar. Drastic increase of smoke levels is observed at the peak load operation in CE at different operating conditions of the vegetable oil, compared with pure diesel operation on CE. This is due to the higher magnitude of the ratio of C/H of CPO (1.13) when compared with pure diesel (0.45). The increase of smoke levels is also due to decrease of air-fuel ratios and VE with vegetable oil compared with pure diesel operation. Smoke levels are related to the density of the fuel. Since vegetable oil has higher density compared to diesel fuels, smoke levels are higher with vegetable oil. However, LHR engine marginally reduced smoke levels due to

efficient combustion and less amount of fuel accumulation on the hot combustion chamber walls of the LHR engine at different operating conditions of the vegetable oil compared with the CE. Density influences the fuel injection system. Decreasing the fuel density tends to increase spray dispersion and spray penetration. Preheating of the vegetable oils reduced smoke levels in both versions of the engine, when compared with normal temperature of the vegetable oil. This is due to i) the reduction of density of the vegetable oils, as density is directly proportional to smoke levels, ii) the reduction of the diffusion combustion proportion in CE with the preheated vegetable oil, iii) the reduction of the viscosity of the vegetable oil, with which the fuel spray does not impinge on the combustion chamber

walls of lower temperatures rather than it directs into the combustion chamber.

Table-5 shows the variation of smoke levels with injection pressure and injection timing at different operating conditions of CPO with different configurations of the engine. Smoke levels decreased with increase of injection timings and

with increase of injection pressure, in both versions of the engine, with different operating conditions of the vegetable oil. This is due to improvement in the fuel spray characteristics at higher injection pressures and increase of air entrainment, at the advanced injection timings, causing lower smoke levels.

TABLE-5

THE VARIATION OF SMOKE INTENSITY AT THE PEAK LOAD OPERATION OILS WITH THE INJECTION TIMING AND INJECTION PRESSURE IN THE CONVENTIONAL AND LHR ENGINES, AT DIFFERENT OPERATING CONDITIONS OF THE VEGETABLE OIL

| Injection timing (°bTDC) | Test Fuel | Smoke intensity (HSU) | | | | | | | | | | | |
|-----------------------------|-----------|---------------------------|-----|-----|----|-----|----|---------------------------|----|-----|-----|-----|----|
| | | Conventional Engine | | | | | | LHR Engine | | | | | |
| | | Injection Pressure (Bars) | | | | | | Injection Pressure (Bars) | | | | | |
| | | 190 | | 230 | | 270 | | 190 | | 230 | | 270 | |
| | | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT |
| 27 | DF | 48 | -- | 38 | -- | 34 | -- | 55 | -- | 50 | -- | 45 | -- |
| | CPO | 70 | 65 | 65 | 60 | 60 | 56 | 65 | 60 | 60 | 55 | 55 | 50 |
| 30 | DF | 36 | -- | 34 | -- | 32 | -- | 45 | -- | 42 | -- | 41 | -- |
| | CPO | 67 | 64 | 64 | 61 | 61 | 58 | 47 | 45 | 45 | 43 | 43 | 41 |
| 31 | DF | 33 | --- | 32 | -- | 30 | -- | 43 | -- | 41 | -- | 40 | -- |
| | CPO | 64 | 61 | 61 | 58 | 58 | 55 | -- | -- | -- | -- | -- | -- |
| 32 | DF | 32 | -- | 31 | -- | 32 | -- | -- | -- | -- | --- | -- | -- |
| | CPO | 60 | 57 | 57 | 54 | 54 | 51 | -- | -- | -- | -- | --- | - |
| 33 | DF | 30 | --- | 30 | -- | 35 | -- | - | -- | -- | -- | -- | -- |

DF-Diesel Fuel, CPO-Crude Pongamia Oil, NT- Normal or Room Temperature , PT- Preheat Temperature

Fig.8 shows the variation of the NO_x levels with BMEP in CE and LHR engine with vegetable oil at the recommended and optimized injection timings at an injection pressure of 190 bar. NO_x levels are lower in CE while they are higher in LHR engine at different operating conditions of the vegetable oil at the peak load when compared with diesel operation. This is due to lower heat release rate because of high duration of combustion causing lower gas temperatures with the vegetable oil operation on CE, which reduced NO_x levels. Increase of combustion temperatures with the faster combustion and improved heat release rates in LHR engine cause higher NO_x levels. As expected, preheating of the vegetable oil further increased NO_x levels in CE and reduced the same in LHR engine when compared with the normal vegetable oil. This is due to improved heat release rates and increased mass burning rate of the fuel leading to increase NO_x emissions in the CE and decrease of combustion temperatures in the LHR engine with the improvement in air-fuel ratios leading to decrease NO_x levels in LHR engine.

Table-6 shows the variation of NO_x levels with injection pressure and injection timing at different operating conditions of CPO with different configurations of the engine. NO_x levels increased with the advancing of the injection timing and with increase of injection pressure in CE with different operating conditions of vegetable oil. Residence time and availability of oxygen had increased, when the injection timing is advanced with the vegetable oil operation, which caused higher NO_x levels in CE. With the increase of injection pressure, fuel droplets penetrate and find oxygen counterpart easily. Turbulence of the fuel spray increased the spread of the droplets which causes the increase the gas temperatures marginally thus leading to increase in NO_x levels as the availability of oxygen and increase of gas temperatures are the two factors responsible for formation of NO_x levels. However, marginal decrease of NO_x levels is observed in LHR engine, due to decrease of combustion temperatures, which is evident from the fact that thermal efficiency is increased in LHR engine due to the reason sensible gas energy is

converted into actual work in LHR engine, when the injection timing is advanced and with increase

of injection pressure.

TABLE-6

THE VARIATION OF NO_x LEVELS AT THE PEAK LOAD WITH THE INJECTION TIMING AND INJECTION PRESSURE IN THE CONVENTIONAL AND LHR ENGINES AT DIFFERENT OPERATING CONDITIONS OF THE VEGETABLE OIL

| Injection timing (° b TDC) | Test Fuel | NO _x levels (ppm) | | | | | | | | | | | |
|----------------------------|-----------|------------------------------|------|------|------|------|------|---------------------------|------|------|------|------|------|
| | | Conventional Engine | | | | | | LHR Engine | | | | | |
| | | Injection Pressure (Bars) | | | | | | Injection Pressure (Bars) | | | | | |
| | | 190 | | 230 | | 270 | | 190 | | 230 | | 270 | |
| | | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT |
| 27 | DF | 850 | ---- | 890 | ---- | 930 | --- | 1300 | -- | 1280 | -- | 1260 | -- |
| | CPO | 740 | 760 | 750 | 770 | 780 | 800 | 1265 | 1250 | 1235 | 1220 | 1200 | 1185 |
| 30 | DF | 935 | --- | 980 | --- | 1020 | -- | 1225 | -- | 1205 | -- | 1185 | -- |
| | CPO | 790 | 805 | 810 | 825 | 820 | 840 | 1190 | 1170 | 1170 | 1140 | 1140 | 1120 |
| 31 | DF | 1020 | --- | 1070 | --- | 1190 | --- | 1150 | -- | 1130 | -- | 1110 | -- |
| | CPO | 850 | 865 | 870 | 885 | 880 | 900 | -- | -- | -- | -- | -- | - |
| 32 | DF | 1105 | ---- | 1150 | --- | 1235 | --- | -- | -- | -- | -- | -- | -- |
| | CPO | 1000 | 1015 | 1020 | 1035 | 1030 | 1050 | -- | - | -- | -- | -- | - |
| 33 | DF | 1190 | ---- | 1230 | --- | 1275 | --- | -- | -- | -- | -- | -- | - |

DF-Diesel Fuel, CPO-Crude Pongamia Oil, NT- Normal or Room Temperature , PT- Preheat Temperature

3.3 COMBUSTION CHARACTERISTICS

Table-7 presents the comparison on the magnitudes of PP, MRPR, TOPP and TOMRPR with the injection timing and injection pressure, at the peak load operation of CE and LHR engine with vegetable oil operation. Peak pressures are lower in CE while they were higher in LHR engine at the recommended injection timing and pressure, when compared with pure diesel operation on CE. This is due to increase of ignition delay, as vegetable oils require large duration of combustion. Mean while the piston started making downward motion thus increasing volume when the combustion takes place in CE. LHR engine increased the mass-burning rate of the fuel in the hot environment leading to produce higher peak pressures. The advantage of using LHR engine for vegetable oil is obvious as it could burn low cetane and high viscous fuels. Peak pressures increased with the increase of injection pressure and with the advancing of the injection timing in both versions of the engine, with the vegetable oil operation. Higher injection pressure produces smaller fuel particles with low surface to volume ratio, giving rise to higher PP. With the advancing of the injection timing to the optimum value with the CE, more amount of the fuel accumulated in the combustion chamber due to increase of ignition delay as the fuel spray found the air at lower pressure and temperature in the combustion

chamber. When the fuel- air mixture burns, it produces more combustion temperatures and pressures due to increase of the mass of the fuel. With LHR engine, peak pressures increases due to effective utilization of the charge with the advancing of the injection timing to the optimum value. The magnitude of TOPP decreased with the advancing of the injection timing and with increase of injection pressure in both versions of the engine, at different operating conditions of vegetable oils. TOPP is more with different operating conditions of vegetable oils in CE, when compared with pure diesel operation on CE. This is due to higher ignition delay with the vegetable oil when compared with pure diesel fuel. This once again established the fact by observing lower peak pressures and higher TOPP, that CE with vegetable oil operation showed the deterioration in the performance when compared with pure diesel operation on CE. Preheating of the vegetable oil showed lower TOPP, compared with vegetable oil at normal temperature. This once again confirmed by observing the lower TOPP and higher PP, the performance of the both versions of the engine is improved with the preheated vegetable oil compared with the normal vegetable oil. This trend of increase of MRPR and decrease of TOMRPR indicated better and faster energy substitution and utilization by vegetable oils, which could replace

100% diesel fuel. However, these combustion characters are within the limits hence the vegetable

oils could be effectively substituted for diesel fuel.

TABLE-7
VARIATION OF PP, MRPR, TOPP AND TOMRPR WITH INJECTION TIMING AND INJECTION PRESSURE AT THE PEAK LOAD OPERATION OF CE AND LHR ENGINE WITH VEGETABLE OIL OPERATION

| Injection timing (°bTDC)/ Test fuel | Engine version | PP(bar) | | | | MRPR (Bar/deg) | | | | TOPP (Deg) | | | | TOMRPR (Deg) | | | |
|---|----------------|--------------------------|------|------|------|--------------------------|-----|-----|-----|--------------------------|----|-----|----|--------------------------|----|-----|----|
| | | Injection pressure (Bar) | | | | Injection pressure (Bar) | | | | Injection pressure (Bar) | | | | Injection pressure (Bar) | | | |
| | | 190 | | 270 | | 190 | | 270 | | 190 | | 270 | | 190 | | 270 | |
| | | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT | NT | PT |
| 27/Diesel | CE | 50.4 | -- | 53.5 | --- | 3.1 | --- | 3.4 | -- | 9 | - | 8 | -- | 0 | 0 | 0 | 0 |
| | LHR | 48.1 | -- | 53.0 | -- | 2.9 | -- | 3.1 | -- | 10 | -- | 9 | -- | 0 | 0 | 0 | 0 |
| 27/ CPO | CE | 45.9 | 47.9 | 48.1 | 49.4 | 2.1 | 2.2 | 2.8 | 2.9 | 12 | 11 | 12 | 10 | 1 | 1 | 1 | 1 |
| | LHR | 58.8 | 59.7 | 62.1 | 63.8 | 3.2 | 3.3 | 3.4 | 3.5 | 11 | 10 | 10 | 9 | 1 | 1 | 1 | 1 |
| 30/CPO | LHR | 60.5 | 61.8 | 64.1 | 64.8 | 3.3 | 3.4 | 3.5 | 3.6 | 10 | 9 | 9 | 9 | 0 | 0 | 0 | 0 |
| 32/CPO | CE | 50.4 | 51.7 | | | 3.0 | 3.1 | | | 11 | 10 | | | 0 | 0 | | |

CPO-Crude Pongmia Oil, CE-Conventional engine, LHR-Low heat rejection, NT-Normal temperature, PT-Preheated temperature,

4. CONCLUSIONS

Vegetable oil operation at 27°bTDC on CE showed the deterioration in the performance, while LHR engine showed improved performance, when compared with pure diesel operation on CE. Preheating of the vegetable oils improved performance when compared with normal vegetable oils in both versions of the engine. Improvement in the performance is observed with the advancing of the injection timing and with the increase of injection pressure with the vegetable oil operation on both versions of the engine. CE with crude vegetable oil operation showed the optimum injection timing at 32°bTDC, while the LHR engine showed the optimum injection at 30° bTDC at an injection pressure of 190 bars. At the recommended injection timing and pressure, crude vegetable oil operation on CE increased smoke levels, decreased NO_x levels, while LHR engine decreased smoke levels and increased NO_x levels when compared with pure diesel operation on CE.

Preheating of the crude vegetable oil decreased smoke levels marginally and increased NO_x levels slightly in CE, while in the LHR engine preheating of the vegetable oils decreased smoke and NO_x levels. CE with vegetable oil operation decreased smoke levels and increased NO_x levels, while LHR engine decreased smoke and NO_x levels with the advancing of the injection timing and increase of injection pressure. Lower peak pressures and more TOPP are observed with normal crude vegetable oil in CE. LHR engine with vegetable oil operation increased PP and decreased TOPP when compared with CE. Preheating increased PP and decreased TOPP when compared with normal vegetable oil operation on both versions of the engine. Lower peak pressures and lower peak gas temperatures are predicted in CE, while higher peak pressures and higher gas temperatures are in the LHR engine with crude vegetable oil operation at the recommended injection timing and pressure.

5. ACKNOWLEDGMENTS

Authors thank authorities of Chaitanya Bharathi Institute of Technology, Hyderabad for providing facilities for carrying out research work. Financial assistance provided by All India Council for

Technical Education (AICTE), New Delhi, is greatly acknowledged.

6. REFERENCES

- [1] Kamo, R., et al., "Injection characteristics that improve performance of ceramic-coated diesel engines". SAE paper No 1999-01-0972, USA, 1999.
- [2] Jaichandar, S. and Tamilporai, P., "Low heat rejection engines - an overview". SAE paper No.2003-01-0405, USA, 2003.
- [3] Ahmaniemi, S. et al.. "Characterization of modified thick thermal barrier coatings", *Journal of Thermal Spray Technology*, Volume-13, No-3, pp:361-369, 2004.
- [4] Ekrem, B., Tahsin, E. and Muhammet, C. Effects of thermal barrier coating on gas emissions and performance of a LHR engine with different injection timings and valve adjustments. *Journal of Energy Conversion and Management* 47, pp. 1298-1310, 2006.
- [5] Parker, D.A. and Dennison, G.M., "The development of an air gap insulated piston", SAE Paper No. 870652, 1987.
- [6] Jabez Dhinagar, S., Nagalingam, B. and Gopala Krishnan, K.V., "A comparative study of the performance of a low heat rejection engine with four different levels of insulation", Proc. of IV International Conference on Small Engines and Fuels, pp: 121-126, Chang Mai, Thailand, 1993.
- [7] Rama Mohan, K., Vara Prasad, C.M. and Murali Krishna, M.V.S., "Performance of a low heat rejection diesel engine with air gap insulated piston", *ASME Journal of Engineering for Gas Turbines and Power*, Volume-121, July, pp: 530-540, 1999.
- [8] Pramanik, K., "Properties and use of jatropha curcas oil and diesel fuel blends in compression ignition engine", *Journal of Renewable Energy*, Vol .28, Issue-2, pp: 239- 48. February 2003.
- [9] Shailendra Sinha and Avinash Kumar Agarawal, "Performance evaluation of a biodiesel (rice bran oil methyl ester) fuelled transportation diesel engine", SAE. Paper No. 2005- 01-1730, 2005.
- [10] Pugazhivadivu, M. and Jayachandran, K., "Investigations on the performance and exhaust emissions of a diesel engine using preheated waste frying oil as fuel", Proc. of the Institution of Mechanical Engineers, *Journal of Automobile Engineering*, 2005.
- [11] Agarwal, A. K., "Bio-fuels (alcohols and biodiesel) applications as fuels for internal combustion engines", *International Journal of Progress in Energy and Combustion Science* 33, pp: 233-271, 2006.
- [12] Gajendra Babu, M.K., Chandan Kumar. and Lalit M. Das., "Experimental investigations on a karanja oil methyl ester fuelled DI diesel engine", SAE. Paper No. 2006-01-0238, 2006.
- [13] Jiwak Suryawanshi, "Performance and emission characteristics of CI engine fueled by coconut oil methyl ester, SAE Paper No. 2006-32-0077, 2006
- [14] Agarwal, D. and Agarwal, A. K. "Performance and emission characteristics of a jatropha oil (preheated and blends) in a direct injection compression ignition engine", *Journal of Applied Thermal Engineering* 27, pp: 2314-2323, 2007.
- [15] Misra, R.D., Murthy, M.S. "Straight vegetable oils usage in a compression ignition engine—A review", *Renewable and Sustainable Energy Reviews*, Volume-14, pp: 3005–3013, 2010.
- [16] Jinlin Xue, Tony E. Grift, Alan C. Hansen, "Effect of biodiesel on engine performances and emissions", *Renewable and Sustainable Energy Reviews* Volume-15, 1098–1116, 2011.
- [17] Bhaskar, T., Nagalingam, B. and Gopala Krishnan, K.V., "The effect of two ignition improving additives

on the performance of jatropha oil in low heat rejection diesel engine” Proceedings of IV International

Conference on Small Engines and their Fuels, pp: 14-19, Thailand, 1993.

[18] Jabez Dhinagar,S., Nagalingam, B.N., Gopalakrishna, K.V., “Experimental investigation of non-edible

vegetable oil operation in a lhr diesel engine for improved performance”, SAE Paper No-932846, 1993.

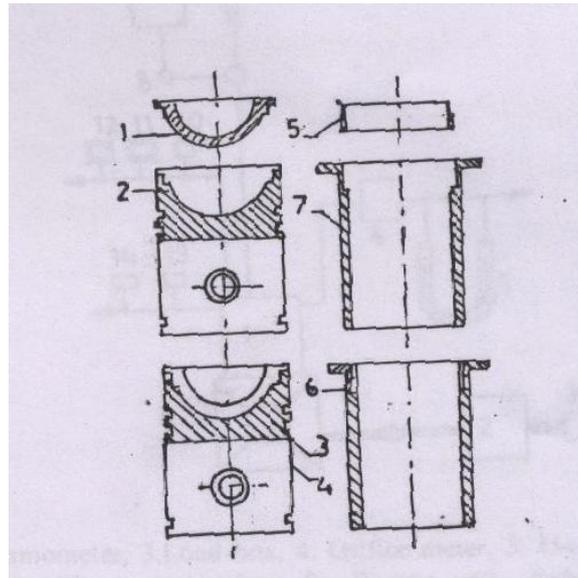
[19] Hanbey Hazar, “Effects of bio-diesel on a low heat loss diesel engine”, Renewable Energy, Volume- 34,

pp:1533–1537, 2009.

[20] Rajendra Prasath, B., P. Tamilporai ,P. and Mohd.Shabir, F., “ Analysis of combustion, performance and

emission characteristics of low heat rejection engine using biodiesel” International Journal of Thermal

Sciences , Volume-49, pp: 2483-2490, 2010



1. Crown

2. Gasket

3. Air gap

4. Body

5 Insert

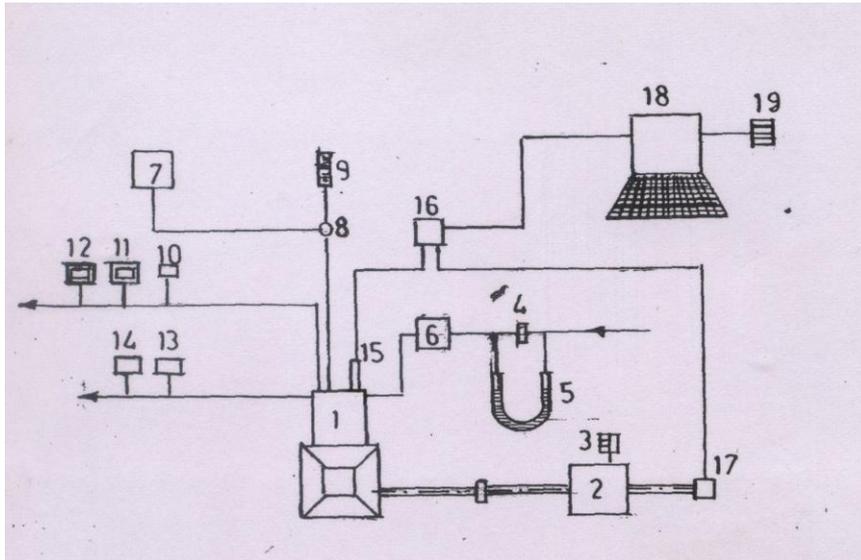
6. Air gap

7. Liner

Insulated piston

Insulated liner

Fig.1 Assembly details of insulated piston and insulated liner



1.Engine, 2.Electical Dynamo meter, 3.Load Box, 4.Orifice meter, 5.U-tube water manometer, 6.Air box, 7.Fuel tank, 8, Pre-heater, 9.Burette, 10. Exhaust gas temperature indicator, 11.AVL Smoke meter, 12.Netel Chromatograph NOx Analyzer, 13.Outlet jacket water temperature indicator, 14. Outlet-jacket water flow meter, 15.Piezo-electric pressure transducer, 16.Console, 17.TDC encoder, 18.Pentium Personal Computer and 19. Printer.

Fig.2 Experimental Set-up

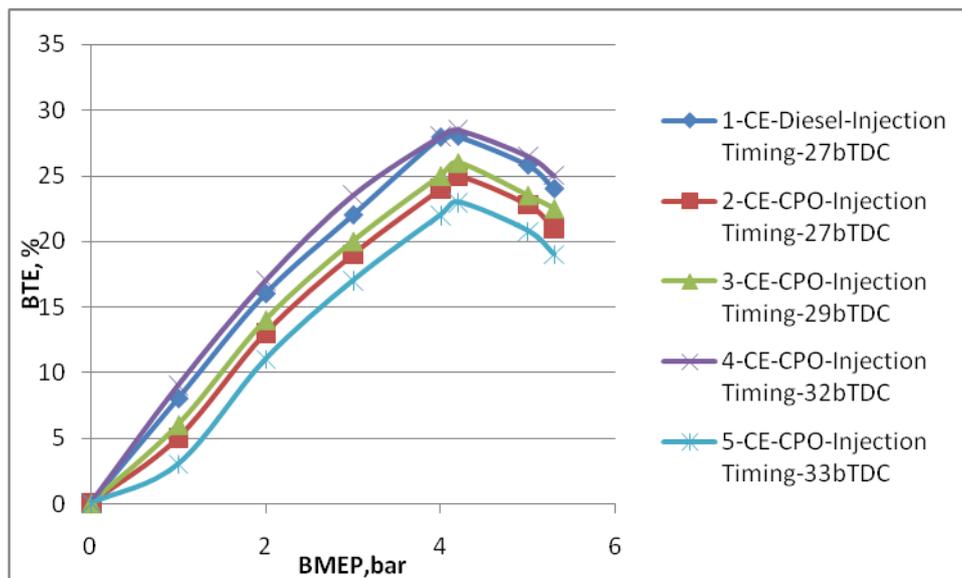


Fig.3 Variation of brake thermal efficiency (BTE) with brake mean effective pressure (BMEP) in conventional engine (CE) at different injection timings with crude pongamia oil (CPO) oil operation.

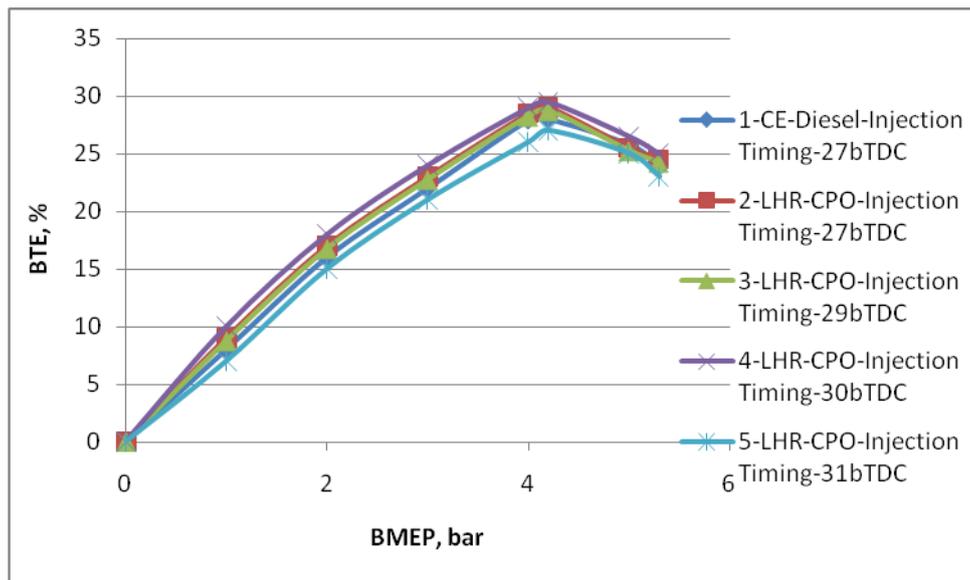


Fig.4 Variation of brake thermal efficiency (BTE) with brake mean effective pressure (BMEP) in low heat rejection (LHR) engine at different injection timings with crude pongamia oil operation (CPO).

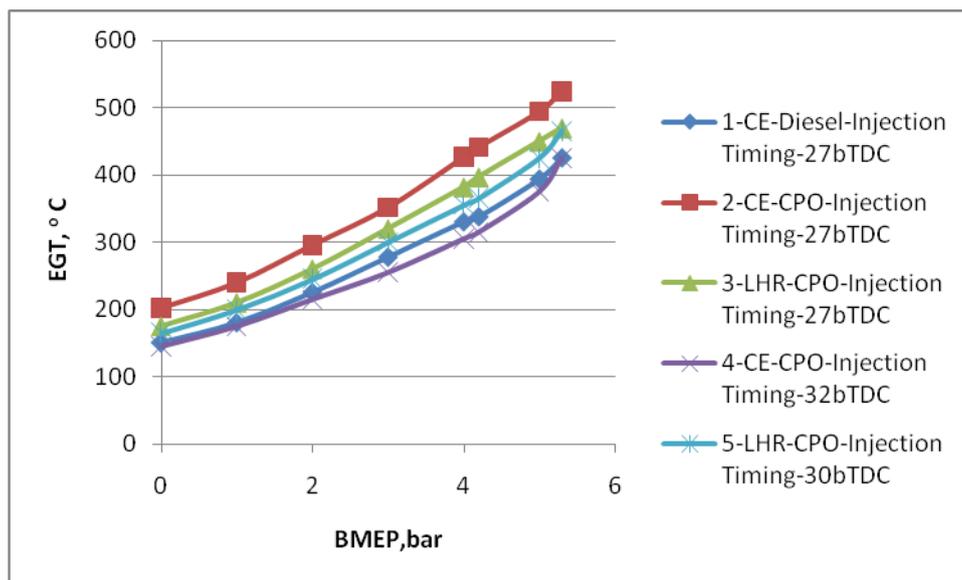


Fig.5 Variation of exhaust gas temperature (EGT) with brake mean effective pressure (BMEP) in conventional engine (CE) and low heat rejection (LHR) engine at recommend injection timing and optimized injection timings with crude pongamia oil (CPO) operation.

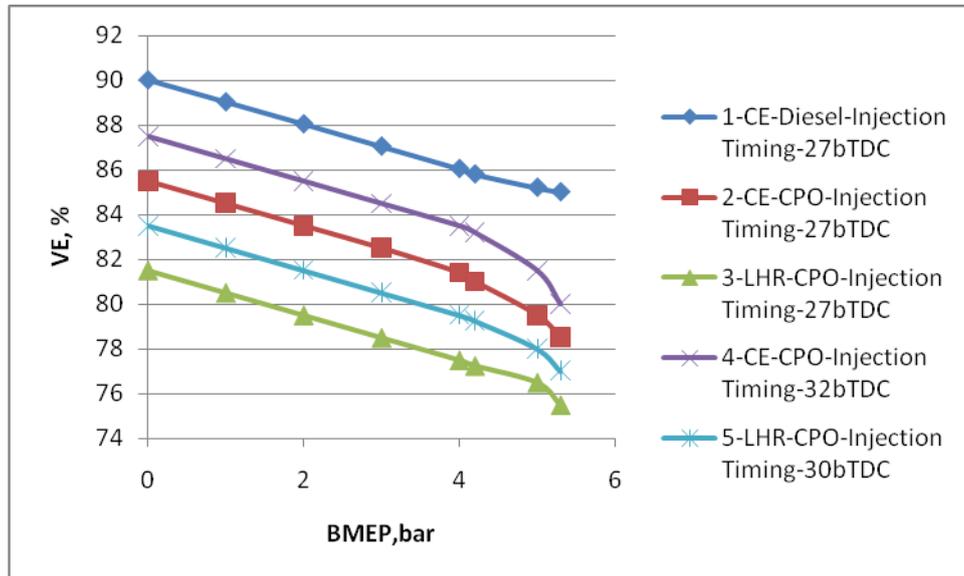


Fig.6. Variation of volumetric efficiency (VE) with brake mean effective pressure (BMEP) in conventional engine (CE) and low heat rejection (LHR) engine at recommend injection timing and optimized injection timings with crude pongamia oil (CPO) operation.

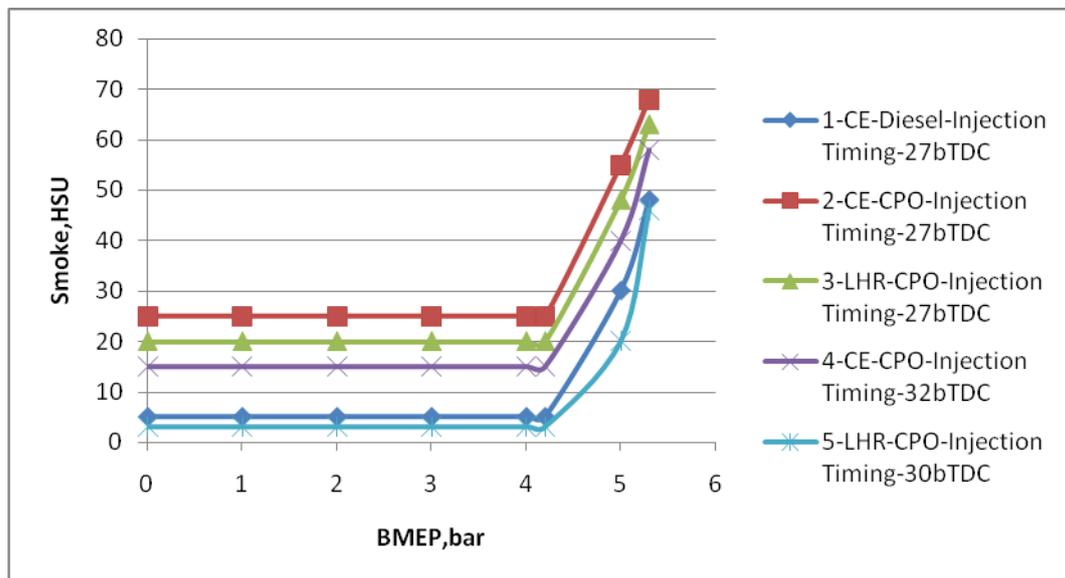


Fig.7. Variation of smoke intensity in Hartridge Smoke Unit (HSU) with brake mean effective pressure (BMEP) in conventional engine (CE) and low heat rejection (LHR) engine at recommend injection timing and optimized injection timings with crude pongamia oil operation (CPO).

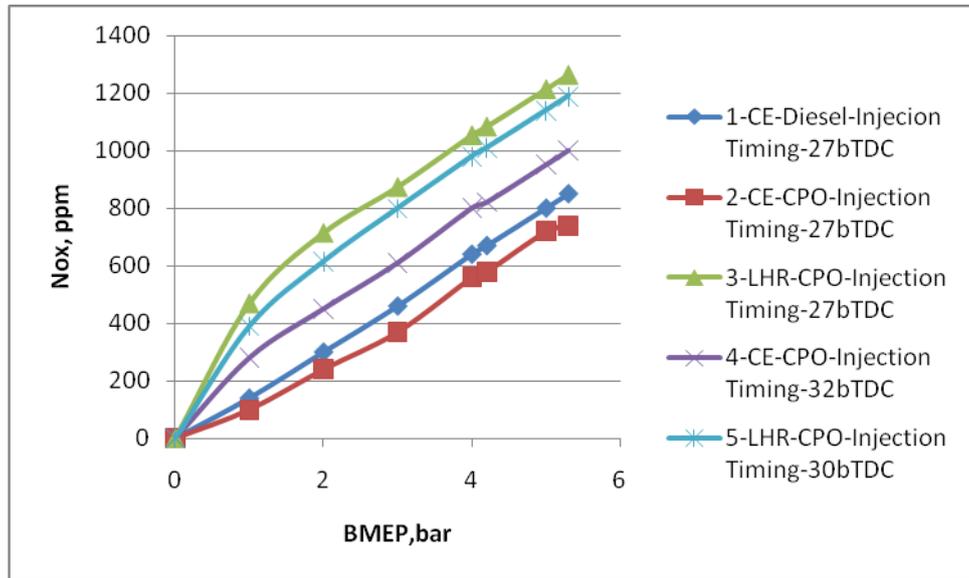


Fig.8. Variation of NOx levels with brake mean effective pressure (BMEP) in conventional engine (CE) and low heat rejection (LHR) engine at recommend injection timing and optimized injection timings with crude pongamia oil (CPO) operation.

Thermal Optimization of Fan assisted Heat Exchanger (Radiator) by Design Improvements

*Prof. D. K. Chavan**, *Prof. Dr. G. S. Tasgaonkar***

**(Department of Mechanical Engineering, MMCOE, Pune-52, University of Pune, India, PhD Scholar, JIT University Rajasthan)*

*** (Department of Mechanical Engineering, ZES, Dnyan Ganga College of Engineering, Narhe - Pune, University of Pune, India)*

ABSTRACT

The Heat Exchangers or Radiators used in automobiles/IC Engines are either rectangular or square in shape, but the air blown/sucked by means of the fan is in circular in area, developing low velocity zones in the corners-hence it is proposed to eliminate corners and develop circular radiators. The object of work is to have a circular radiator which is compact-made with minimum material-less costly-more efficient-that will work with minimum power consumption of fan and maximum utilization of air flow.

It is proposed to develop three different types of radiators and the results of one rectangular and two circular radiators will be compared. Here results like velocity, flow rate of water, and temperature at different points of the radiator are compared. After validating the present concept through actual manufacturing and mathematical calculations it can be commercially applied to applications such as automobile radiators, IC Engine radiators, heat exchangers used in refrigerators and air-conditioners etc.

Considering the number of vehicles, refrigerators, air-conditioners used at National and International level, a slight modification/improvement in efficiency and reduction in cost will add to the economy a great extent.

Keywords – Circular radiator, smaller fins, thermal optimization, velocity, validation etc

I. INTRODUCTION

The present manufacturing of heat exchangers commonly used in practice in Automobiles, Internal Combustion (IC) engines, Refrigeration systems, and Power plants emphasizes on production process, materials and spacing of fins.[1]

The proposed work aims at optimizing the fan assisted heat exchanger (radiator) by improvement in the design.

II. ORIGIN OF THE RESEARCH PROBLEM

The present heat exchangers/radiators are rectangular in shape. But the air blown by the fan is circular in area, developing low velocity area in the corners.[3] Therefore circular radiators which are compact are proposed to be developed and tested to improve the efficiency.

No significant work has been done in this area.

III. INDENTATIONS AND EQUATIONS

Figure1 and 2 shows a rectangular/square-shaped heat exchanger with a fan provided to deliver air in a circular area.

If the length and breadth of the heat exchanger is equal to D , the effective area of such heat exchanger will be equal to D^2 .

While the flow of air from the fan (without shroud) will be of area $(\pi/4) D^2 = 0.76 D^2$.

The difference in the area of the square and the circle would be $\{D^2 - (\pi/4) D^2\} = 0.24 D^2$. [a]

IV. FIGURES

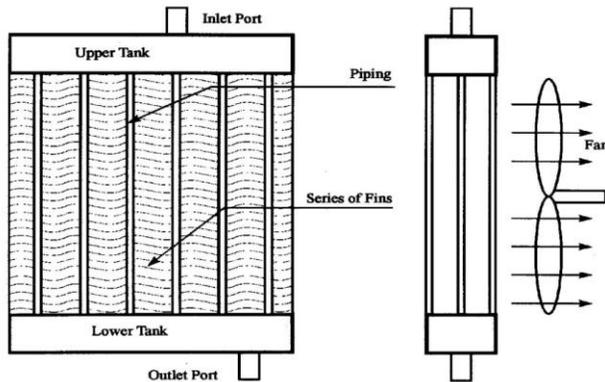


Fig 1: Existing rectangular radiator

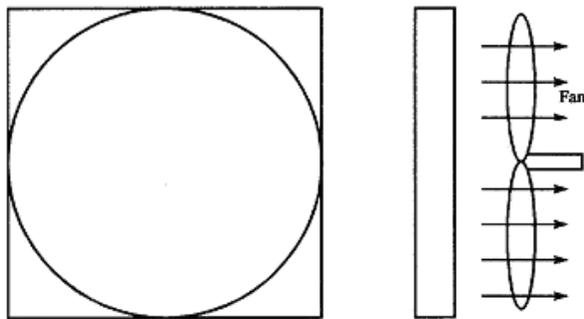


Fig 2: Air-cooled Square shaped heat exchanger

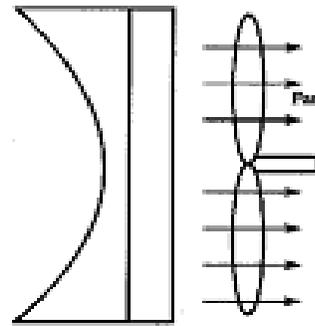


Fig 3: Velocity profile of an air-cooled heat exchanger using a fan

As shown in Fig (3) velocity of air generated by the fan is not constant along its axial direction. It is found to be almost zero at the centre and gradually increases at the rate of square of the radius towards the periphery [4].

V. FIN GEOMETRY FOR PROPOSED HEAT EXCHANGER

Since the velocity of air varies from centre outwards, smaller fins are provided at the centre and longer fins at the periphery. The length of fins are so adjusted that the velocity of air coming out of the heat exchanger remains constant over the effective area. (fig.4)

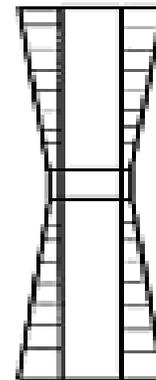


Fig 4: Fin geometry for proposed Heat Exchanger

It is proposed to develop 3 different types of radiators, one is almost square and two are circular.

- 1) First one is almost square shaped.
- 2) Circular first type –in this arrangement distance between the tubes is same but the fin size increases towards periphery.[5]
- 3) Circular second type –in this arrangement distance between the tubes decreases-i.e., more number of tubes are provided at the periphery and fin size is same.

All the 3 radiators will be manufactured out of,

- a) Same material
- b) Same tube size
- c) Same fin thickness

Fig. 5 shows the proposed circular radiator.

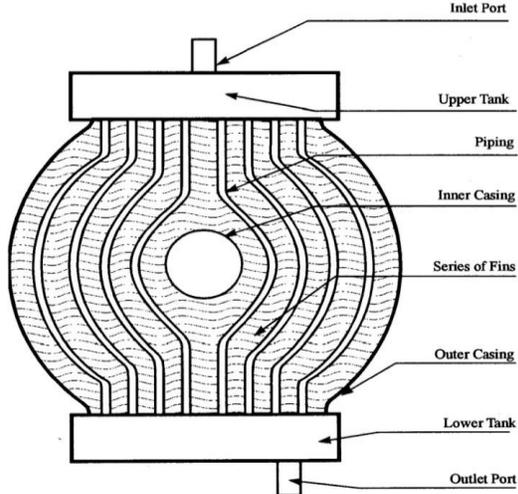


Fig 5: Proposed Heat Exchanger (Radiator)

VI. PROPOSED TEST SET UP

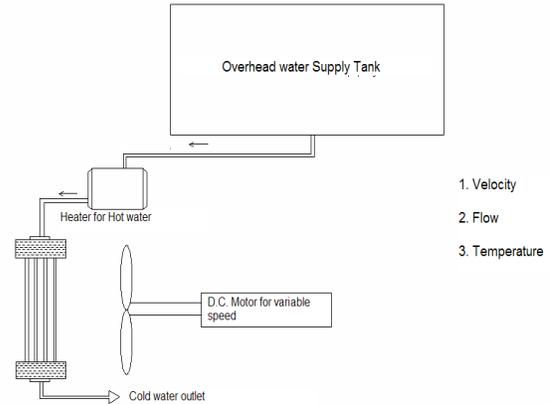


Fig 6: Proposed Test set up

It consists of,

- 1) Heater-to supply hot water to @ 70-80 °C.
- 2) Radiator –which can be changed.
- 3) D.C. motor for variable speed.
- 4) Velocity of air at different points is measured with Anemometer.
- 5) Flow rate of water is measured with Rotameter.
- 6) Temperature at various points is measured with Infrared Temperature measuring gun [b,c]

Journal Papers:

1. ASME 2002 International Mechanical Engineering Congress & Exposition Heat Transfer Vol 6, Nov 17-22, 2002 USA Paper No. IMECE 2002-33888
2. Optimization of Central heating radiator by Cihat Arslanturk & A. Feridum Ozguc. Applied Engg Vol 83, issue 11, Nov 2006, Pages 1190-1197.
3. Shape optimization of vehicle radiator using CFD by, Maddiptla, Sridhar, Guessous, Laila, American Physical Society, Division

of Fluid Dynamics, 55th Annual meeting
KA006M dated 11/2002.

4. Parametric studies on Automotive Radiators
by, C. Oliet, A. Oliva, J.Castro & C.D.
Perege- Segarre; Applied Thermal Engg
Journal vol 27, issue 11-12;Aug 2007, Pages
2033-2043.
5. Designing a more effective car radiator
:Application note Maple soft, Maple Inc
2008
6. The Patent search carried by Patent attorney.

VII. KINDS CONCLUSION AND THEIR EXPECTED RESULTS

The new proposed heat exchanger will be,

1. Circular in shape
2. Compact
3. Less material requirement
4. Less power consumption for fan
5. More efficient
6. Since material saving is @ 24%, cost saving
on mass scale production will be @ 20%
once the dies are manufactured.

Books:

- a) Heat Exchanger Design (II Edition) Arthur P
Fraas. Published by John Wiley & Sons
New York
- b) Heat Exchangers, Selection, rating, &
Thermal design (II edition) Sadik Kakak &
Hongtan Lin
- c) Fundamentals of Heat Exchanger Design
Ramesh K Shah &Dusan P Sekulic
- d) Heat Exchanger Design Handbook by
T Kuppan

SIMULTANEOUS TUNING OF POWER SYSTEM STABILIZER PARAMETERS FOR MULTIMACHINE SYSTEM

Mr.M.Sivasubramanian¹
Assistant Professor / EEE

Mr.P.Musthafa²
Assistant Professor / EEE

Mr.K Sudheer³
Assistant Professor / EEE

VELTECH MULTITECH Dr RANGARAJAN Dr SAKUNTHALA ENGINEERING COLLEGE, AVADI,
CHENNAI, INDIA

ABSTRACT: *Optimal multiobjective design of robust multimachine power system stabilizers (PSSs) using genetic algorithms is presented in this thesis. A conventional speed-based lead-lag PSS is used. The multimachine power system operating at various loading conditions and system configurations is treated as a finite set of plants. The stabilizers are tuned to simultaneously shift the lightly damped and undamped electromechanical modes of all plants to a prescribed zone in the s-plane. A multiobjective problem is formulated to optimize a composite set of objective functions comprising the damping factor, and the damping ratio of the lightly damped electromechanical modes.*

The problem of robustly selecting the parameters of the power system stabilizers is converted to an optimization problem which is solved by a genetic algorithm with the eigenvalue-based multiobjective function. The effectiveness of the suggested technique in damping local and interarea modes of oscillations in multimachine power systems, over a wide range of loading conditions and system configurations, is confirmed through eigenvalue analysis and nonlinear simulation results

Index Terms - *Small signal stability, genetic algorithms, multiple objective optimization, robustness, simultaneous stabilization.*

1. INTRODUCTION

In multi-machine power systems with several poorly damped modes of oscillations, several power system stabilizers (PSS) need to be on-line and optimally tuned. With present-day large-scale systems comprising many interconnected machines, the problem of PSS tuning is not a straight-forward exercise, and in

some cases can become relatively too complex to resolve. The problem of PSS tuning is further complicated by the fact that operating conditions in a power system are continuously varying. Therefore tuning the PSS such that, it would provide a satisfactory performance over the entire range of variations is a rather exhaustive exercise.

Research has been directed towards the design of adaptive (self-tuning), variable structure and other control strategies that provide robust tuning. However, implementation of such PSS requires continuous on-line calculation of an identified model using parameter estimation and evaluation of the control strategy. In recent years, research has been directed towards the application of advanced numerical computation methods such as neural networks and genetic algorithms (GA) to PSS tuning.

This paper presents the design of a GA based PSS that uses an eigen value based parameter optimization criterion to determine the *fitness function* of an individual within a population of possible solutions. Genetic algorithms are global search techniques and provide a powerful tool for optimization problems by miming the mechanisms of natural selection and genetics. These operate on a population of potential solutions applying the principle of survival of the fittest to produce better and better approximations to a solution. In each generation, a new set of approximations is created by the process of selecting the individuals according to their level of fitness in the problem domain and breeding them together using operators borrowed from natural genetics [1]. Thus, the population of solutions is successively improved with respect to the search objective, by replacing least fit individuals with new ones (offset of individuals from the previous generation), better suited to the environment, just as in natural evolution. The performance (*fitness*) of each

individual in the problem domain is assessed through an objective function that ultimately establishes the basis for the biased selection process. Higher the individuals fitness is, higher is its chance to pass on genetic information to successive generations. The selected individuals are then modified through the application of *genetic operators*, in order to obtain the next generation. Thus GA based optimization of PSS parameters is more likely to converge to the global optima than a conventional optimization, since they search from a population of possible solutions, and are based on probabilistic transition rules. Moreover, by tuning the PSS simultaneously, the eigenvalue drift problem is eliminated. In the recent literature, application of genetic algorithm to tune the parameters of PSS has been reported [1], [2], [7]. A GA based optimization method has been used in [2] to tune the parameters of a rule-based PSS. This way, the advantages of the rule-based PSS such as its robustness, less computational burden and ease of realization are maintained. Introduction of GA helps obtain an optimal tuning for all PSS parameters simultaneously, which thereby takes care of interactions between different PSS. In [2] simultaneous tuning for all the PSS in the system using a GA based approach has been developed. The GA seeks to shift all eigenvalues of the system within a region in the stable domain.

2. APPLICATION OF GENETIC ALGORITHM TO PSS DESIGN

Each individual in the initial population has an associated objective function value. Using the objective function information, the GA then produces a new population. The application of a genetic algorithm involves repetitively performing two steps:

1. The calculation of the objective functions for each of the individuals in the current population. To do this, the system eigenvalues must be computed.
2. The genetic algorithm then produces the next generation of individuals using the selection, crossover and mutation operators.

These two steps are repeated from generation to generation until the population has converged, producing the optimum parameters. A genetic algorithm (GA)-based approach to robust PSS design, in which several operating conditions and system configurations are simultaneously considered in the design process, is presented. The advantage of the GA technique is that it is independent of the complexity of the performance

index considered. It suffices to specify the objective function and to place finite bounds on the optimized parameters. Initially, the robust PSS design was formulated as a single objective function problem, and not all PSS parameters were considered adjustable. However, in practice, one is typically confronted with multiple objective functions and these objective functions are generally of diverse natures. In this thesis, the problem of robust PSS design is formulated as a multiobjective optimization problem and GA is employed to solve this problem. Moreover, unlike [2], all PSS parameters were considered adjustable, and more severe disturbances were used to assess the potential of the multiobjective approach. Robustness is achieved by considering several operating conditions and system configurations simultaneously.

The multiobjective problem is concocted to optimize a composite set of two eigenvalue-based objective functions comprising the desired damping factor, and the desired damping ratio of the lightly damped and undamped electromechanical modes. The use of the first objective function will result in PSSs that shift the lightly damped and undamped electromechanical modes to the left-hand side of a vertical line in the complex s -plane; hence, improving the damping factor. The use of the second objective function will yield PSSs' settings that place these modes in a wedge-shape sector in the complex s -plane, thus improving the damping ratio of these modes. Consequently, the use of the multiobjective function will therefore guarantee that the relative stability and the time domain specifications are concurrently secured. The proposed design approach has been applied to a multimachine power system. The eigenvalue analysis and the nonlinear simulation results have been carried out to assess the effectiveness of the proposed PSSs under different disturbances, loading conditions, and system configurations.

3. CONTROLLER TUNING:

The problem of selecting the parameters of the controllers that would assure maximum damping performance over the considered set of operating points is solved via a GAs optimization procedure with an eigenvalue based performance index.

A. MODEL AND CONTROL STRUCTURE

Equations 1 describe a linear model of power system extracted around a certain operating point.

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \quad (1)$$

The controller is a lead-lag type described by:

$$V(s) = K(s)y(s) \quad (2)$$

where $K(s)$ is the transfer function of the controller, $y(s)$ is the measurement signal and $V(s)$ is the output signal from the controller which will provide additional damping by moving modes to the left. Equation 2 can be expressed in the statespace form as:

$$\begin{aligned} \dot{x}_k &= A_k x_k + B_k y \\ u &= C_k x_k + D_k y \end{aligned} \quad (3)$$

where x_k is the state vector of the controller. Combining Equations 1 and 3 with Equations 1 and 2 a closed loop system given in Equation 4 is obtained.

$$\dot{x}_{cl} = A_{cl} x_{cl} \quad (4)$$

Let $\lambda_i = \sigma_i \pm j\omega_i$ be the i -th eigenvalue (mode) of the closed loop matrix. Then, the damping coefficient (ζ) of the i -th eigenvalue is defined by

$$\zeta_i = -\frac{\sigma_i}{\sqrt{\sigma_i^2 + \omega_i^2}} \quad (5)$$

The structure of PSS is given below.

$$U(s) = K_i \frac{sT_{wi}}{1+sT_{wi}} \left[\frac{(1+sT_{1i})(1+sT_{3i})}{(1+sT_{2i})(1+sT_{4i})} \right] \Delta\omega_i(s)$$

B. OBJECTIVE FUNCTION

Very often, the closed loop modes are specified to have some degree of relative stability. In this case, the closed-loop eigenvalues are constrained to lie to the left of a vertical line corresponding to a specified damping factor. Select the parameters of the PSS to minimize the following objective function:

$$J_1 = \sum_{j=1}^{np} \sum_{\sigma_{i,j} \geq \sigma_0} [\sigma_0 - \sigma_{i,j}]^2 \quad (6)$$

where np is the number of operating points considered in the design process, and $\sigma_{i,j}$ is the real part of the i -th eigenvalue of the j -th operating

point. The relative stability is determined by the value of σ_0 .

In many cases, certain time-domain control system specifications such as maximum overshoot, rise time and steady-state error goals can be realized by placing the closed-loop eigenvalues of the system within a region bounded by minimum of the damping coefficients in the left-half of the complex s -plane. In order to do this, the objective function of (7) is changed to:

$$J_2 = \sum_{j=1}^{np} \sum_{\zeta_{i,j} \leq \zeta_0} [\zeta_0 - \zeta_{i,j}]^2 \quad (7)$$

where $\zeta_{i,j} \leq \zeta_0$ is the damping ratio of the i -th eigenvalue of the j -th operating point. This will place the closed-loop eigenvalues in a wedge-shaped sector in which $\zeta_{i,j} \geq \zeta_0$ as shown in Fig. 4.2

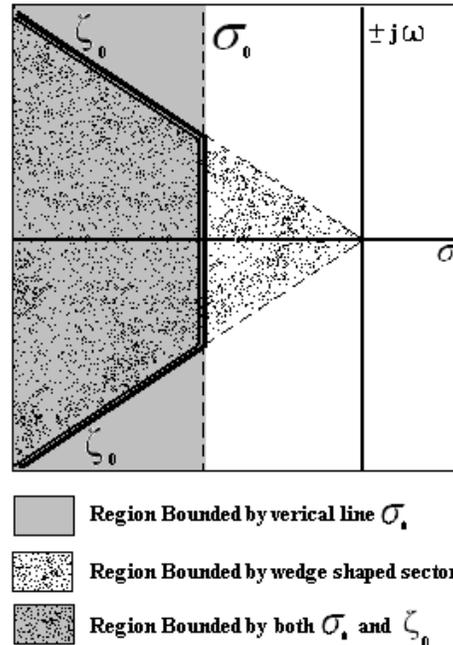


Fig.1 Interested region of pole locations in s-plane

These single objective problems may be converted to a multiple objective problem by assigning distinct weights to each objective. In this case, the conditions $\sigma_{i,j} \leq \sigma_0$ and $\zeta_{i,j} \geq \zeta_0$ are imposed simultaneously. The parameters of the PSS may be selected to minimize the following objective function:

$$J = J_1 + aJ_2 \quad (8)$$

$$J = \sum_{j=1}^{np} \sum_{\sigma_{i,j} \geq \sigma_0} [\sigma_0 - \sigma_{i,j}]^2 + a \sum_{j=1}^{np} \sum_{\zeta_{i,j} \leq \zeta_0} [\zeta_0 - \zeta_{i,j}]^2$$

This will place the system closed-loop eigenvalues in the D-shape sector as shown in Fig.4.3. The design problem can be formulated as the following constrained optimization problem, where the constraints are the PSS parameter bounds:

Minimize J subject to

$$\left\{ \begin{array}{l} K_{i,\min} \leq K_i \leq K_{i,\max} \\ T_{1i,\min} \leq T_{1i} \leq T_{1i,\max} \\ T_{2i,\min} \leq T_{2i} \leq T_{2i,\max} \\ T_{3i,\min} \leq T_{3i} \leq T_{3i,\max} \\ T_{4i,\min} \leq T_{4i} \leq T_{4i,\max} \end{array} \right. \quad (9)$$

The minimization of the objective function J will result in a PSS structure that satisfies the time-domain performance specifications as well as relative stability. It is necessary to mention here that if only particular eigenvalues need to be relocated, then only those eigenvalues should be taken into consideration in the computation of the objective function. This is usually the case in dynamic stability where it is desired to relocate the electromechanical modes of oscillations. The proposed approach employs GA to solve this optimization problem and search for optimal or near optimal set of PSS parameters, $K_i, T_{1i}, T_{2i}, T_{3i}, T_{4i}$ for $i=1$ to m , where m is the number of machines.

C. CONTROL PARAMETERS AND GA PARAMETERS

A two stage lead/lag compensator structure was chosen for the PSS. Hence all the five parameters (K, T_1, T_2, T_3 and T_4) are taken as control parameters. Gain K is bounded between 0.5 and 150 and all Time constants are bounded between 0.01 and 1.5 seconds. These parameter-bounds were defined on the basis of conventional control design for nominal operating condition. In GA implementation, the crossover and mutation probabilities of 0.95 and 0.033, respectively, are found to be quite satisfactory. The number of individuals in each generation is selected to be 200. In addition, the search will terminate if the best solution does not change for more than 50 generations or the number of generations reaches 100 for single objective function and 200 for multi-objective function respectively.

4. TEST SYSTEM AND PSS DESIGN

The one-line diagram of the test system is given in Fig. 2. This two-area power system, which as a benchmark system for inter-area oscillations studies consist of two generators in each area, connected via a 220 km tie line. All generators are equipped with simple exciters and have the same parameters. Damping control is provided at all four generators.

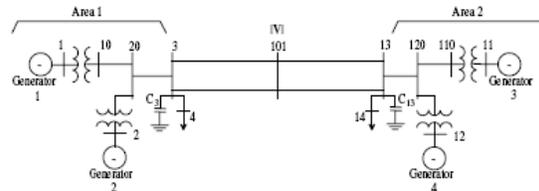


Fig 2. Two Area 4 Machine Power System

To design the proposed PSSs, two different operating conditions that represent the system under severe loading conditions and critical line outages in addition to the base case are considered. These conditions are extremely harsh from the stability viewpoint [6]. Two system configurations, which is heavily loaded with 400 MW of power flowing from area 1 to area 2, were analyzed:

- Operating Condition 1 — System with two lines between bus 3 and 101
- Operating Condition 2 — System with a single line between bus 3 and 101

There are 30 parameters to be optimized, namely $K_i, T_{1i}, T_{2i}, T_{3i}, T_{4i}$ $i=1,2,3$. The time constant T_w is set to be 5 s [7]. In this study, σ_0 and ζ_0 are chosen to be 1.0 and 0.20, respectively.

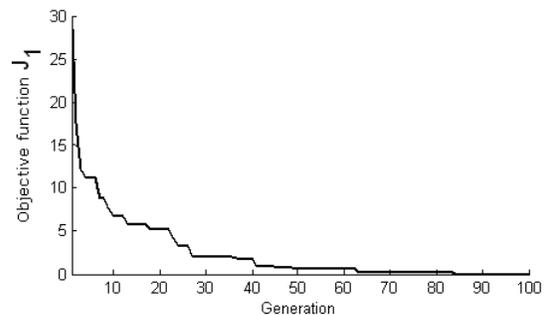


Fig. 3 Convergence for objective function J_1

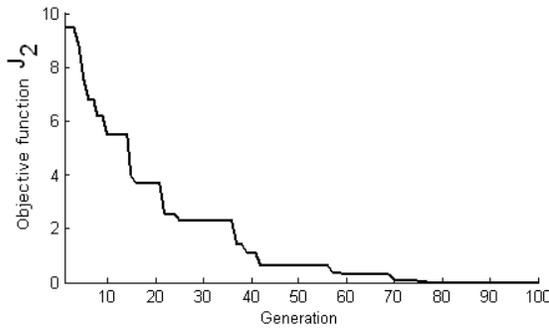


Fig. 4 Convergence for objective function J_2

Several values for the weight a were tested; it was found that the effect of varying a on the final goals is minimal. The results presented here are for $a=10$. The convergence rate of the single objective functions and, and the multiobjective function are shown in Fig. 4, 5 and 6. The final value of the objective functions J_1 and J_2 is 0, indicating that all of the electromechanical modes have been shifted to the left of the vertical line $\sigma_0 = -1$ and $\zeta_0=0.2$ respectively. The final value of the objective function $J = J_1 + aJ_2$ is $J=0$, indicating that all of the electromechanical modes have been shifted to the specified D-shape sector in the s-plane.

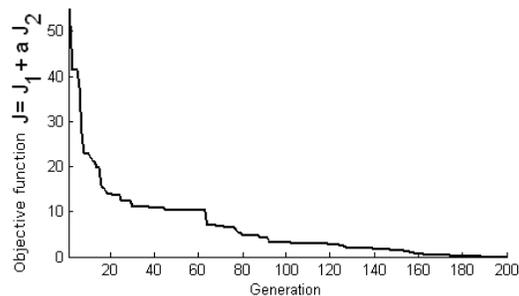


Fig. 5 Convergence for objective function J

Table I Tuned PSS parameters for Objective function J_1 , J_2 and J .

| Obj | Gen | K | T_1 | T_2 | T_3 | T_4 |
|-------|-----|---------|--------|--------|--------|--------|
| J_1 | G1 | 39.2777 | 0.7111 | 0.0359 | 0.274 | 0.0695 |
| | G2 | 42.6168 | 1.127 | 0.0476 | 0.359 | 0.0561 |
| | G3 | 20.1555 | 0.774 | 0.0287 | 0.630 | 0.0685 |
| | G4 | 34.5081 | 0.1737 | 0.0617 | 0.2445 | 0.0714 |
| J_2 | G1 | 38.9357 | 0.8276 | 0.0247 | 0.7307 | 0.0555 |
| | G2 | 31.7945 | 0.9154 | 0.0383 | 0.8157 | 0.0397 |
| | G3 | 34.2916 | 0.7733 | 0.0248 | 1.095 | 0.0479 |
| | G4 | 10.2385 | 0.1612 | 0.0953 | 1.1954 | 0.0466 |
| J | G1 | 48.8622 | 0.3686 | 0.0137 | 0.445 | 0.0159 |
| | G2 | 28.6638 | 0.7259 | 0.0252 | 0.6528 | 0.037 |
| | G3 | 42.938 | 0.7016 | 0.0426 | 0.5638 | 0.0403 |
| | G4 | 49.4392 | 0.1211 | 0.0619 | 0.3043 | 0.0228 |

5. SMALL SIGNAL AND LARGE SIGNAL TESTS

Small-signal analysis provides a mean to compare the damping of the different system modes.

Table II. Eigenvalues and Damping ratios of electromechanical modes with and without PSSs

| | Case K2L | | | Case K1L | | |
|---------------|----------|---------------|---------|----------|---------------|---------|
| | σ | $\pm j\omega$ | ζ | σ | $\pm j\omega$ | ζ |
| With out PSSs | 0.191 | 5.808 | -0.03 | 0.195 | 5.716 | -0.03 |
| | 0.088 | 4.002 | -0.02 | 0.121 | 3.798 | -0.03 |
| | -0.028 | 9.649 | 0.003 | 0.097 | 6.006 | -0.01 |
| With PSSs | -1.198 | 12.649 | 0.094 | -1.26 | 12.157 | 0.103 |
| | -1.276 | 11.827 | 0.107 | -1.24 | 11.799 | 0.105 |
| | -1.080 | 10.782 | 0.100 | -1.05 | 10.784 | 0.098 |
| J_2 | -2.887 | 12.496 | 0.225 | -3.06 | 12.561 | 0.237 |
| | -3.543 | 11.319 | 0.299 | -3.47 | 11.228 | 0.295 |
| | -2.894 | 10.996 | 0.255 | -2.78 | 10.961 | 0.247 |
| With PSSs | -3.281 | 14.606 | 0.219 | -3.27 | 14.494 | 0.220 |
| | -2.739 | 13.119 | 0.204 | -2.61 | 12.395 | 0.206 |
| | -2.632 | 11.242 | 0.228 | -2.64 | 11.083 | 0.232 |

To better understand the results, we have completed the small signal analysis of the tuned PSS was performed on the system for both single tie-line (K1L) and on two tie-lines (K2L) configurations. The system electromechanical modes, for the base case and the two operating conditions (cases K1L–K2L), without and with the PSSs tuned using J_1 , J_2 and J are listed in Table II.

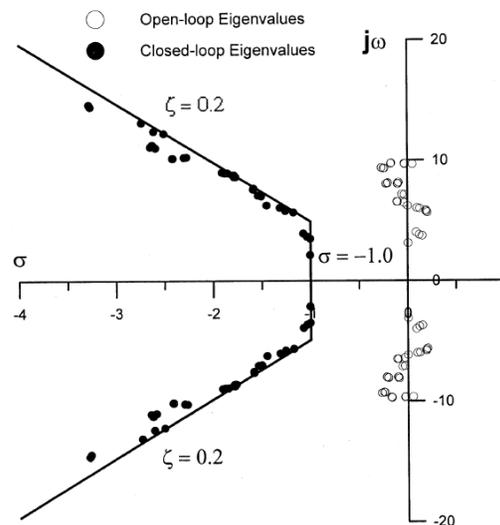


Fig. 6 Eigenvalues associated with modes of J

In assessing PSS, small-signal performance is not enough. Good performance during large perturbations and good robustness with respect to changing operating conditions are

other criteria of an equal importance. To demonstrate the effectiveness of the PSSs tuned using the proposed multiobjective function over a wide range of operating conditions, the following disturbance is considered for nonlinear time simulations.

Table III. Test Cases for Large Signal Assessment

| TEST CASE NAME | SYSTEM CONFIGURATION | CONTINGENCY DESCRIPTION |
|----------------|----------------------|---|
| A | K2L SYSTEM | 5 cycle, three phase fault at bus 101 with the outage of 230 KV line. |
| B | K2L SYSTEM | 5 cycle, three phase fault at bus 1 with no equipment outage. |
| C | K1L SYSTEM | 3 cycle, single phase fault at bus 120 without outage. |

It is clear that the system response from Fig.7 that, the PSSs tuned using the multiobjective function J settles faster, and provides superior damping in comparison with the case when either of J_1 or J_2 is used. This indicates that the time domain specifications were simultaneously met.

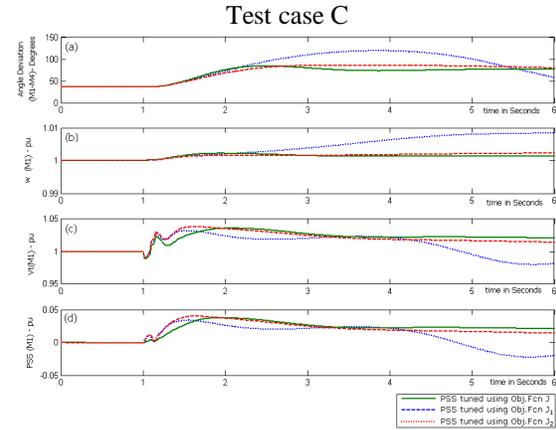
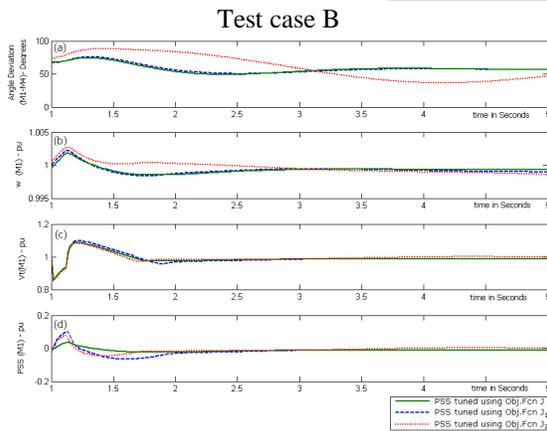
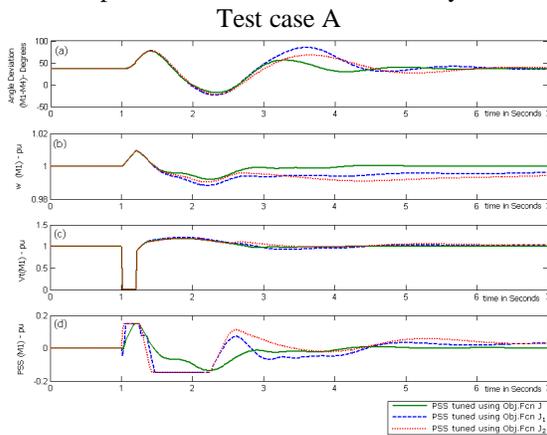


Fig. 7 Responses for Test case A, B and C respectively.

- (a) Angular Deviation (δ_{14}) between (M1-M4) in Degrees
- (b) Speed (ω_1) of M1 in pu.
- (c) Terminal voltage (V_t) of M1 in pu
- (d) PSS output for M1 in pu.

7. CONCLUSION

In this thesis, optimal multiobjective design of robust multimachine power system stabilizers (PSSs) using GAs is proposed. A conventional speed-based lead-lag PSS is used in this work. The multimachine power system operating at various loading conditions and system configurations is treated as a finite set of plants. The stabilizers are tuned to simultaneously shift the lightly damped electromechanical modes of all plants to a prescribed zone in the s-plane. A multiobjective problem is formulated to optimize a composite set of objective functions comprising the damping factor, and the damping ratio of the lightly damped electromechanical modes. The problem of robustly selecting the parameters of the power system stabilizers is converted to an optimization problem which is solved by a GA with the eigenvalue-based multiobjective function. The eigenvalue analysis and non linear time-domain simulations, confirms that the closed-loop plant performance is consistent with the design requirements in spite of changes in the operating conditions, and reveals the superiority of the PSSs tuned using the multiobjective function in damping local and inter-area modes of oscillations.

REFERENCES

- [1] J. J. Grefenstette, "Optimization of control parameters for genetics algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, pp. 122–128, Jan./Feb. 1986.
- [2] Y. L. Abdel-Magid, M. A. Abido, S. Al-Baiyat, and A. H. Mantawy, "Simultaneous stabilization of multimachine power systems via genetic algorithms," *IEEE Trans. Power Syst.*, vol. 14, pp. 1428–1439, Nov. 1999.
- [3] E. Larsen and D. Swann, "Applying power system stabilizers," *IEEE Trans. Power Appar. Syst.*, vol. PAS-100, pp. 3017–3046, 1981.
- [5] C. M. Lim and S. Elangovan, "Design of stabilizers in multimachine power systems," *Proc. Inst. Elect. Eng.*, pt. C, vol. 132, no. 3, pp. 146–153, 1985.
- [6] P. Kundur, M. Klein, G. J. Rogers, and M. S. Zywno, "Application of power system stabilizers for enhancement of overall system stability," *IEEE Trans. Power Syst.*, vol. 4, pp. 614–626, May 1989.
- [7] M. J. Gibbard, "Robust design of fixed-parameter power system stabilizers over a wide range of operating conditions," *IEEE Trans. Power Syst.*, vol. 6, pp. 794–800, May 1991.
- [8] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

Transmission Loss Allocation and Loss Minimization By Incorporating UPFC in LFA

¹Sunil Kumar A V, ²Roopa V, ³Javid Akthar, ⁴Dr. Shivasharanappa G C

¹(Electrical & Electronics dept GCE, Ramanagaram Karnataka) INDIA

²(Electrical & Electronics dept, GCE, Ramanagaram, Karnataka) INDIA

³(Electrical & Electronics dept GCE, Ramanagaram Karnataka) INDIA

⁴(Electrical & Electronics dept DBIT, Bangalore ,Karnataka) INDIA.

Abstract-The paper focuses on the issue of transmission loss allocation and transmission loss minimization by incorporating UPFC injection model using load flow analysis. To investigate the effect of the UPFC on the steady state condition of the system and load flow, different models can be used. These models are usually based on modification of traditional load flow methods. In this paper, a mathematical model for UPFC extended to UPFC injection model. Since accurate power tracing is very difficult as well as an allocation of losses for a particular transaction (in power business it is buying and selling system) may not be effectively realized. However, loss allocation is an important aspect in determining the cost of transmission. Thus a methodology to find the losses accurately is vital. It is imperative to make sure that all users of the transmission network are charged proportionate to their usage and this aspect is more important because of the common infrastructure. The Z-bus loss allocation method is used to achieve the required objective. This method will promote more efficient network operations when implemented in deregulated electric industries. The Unified Power Flow Controller (UPFC) injection model is incorporated in Load Flow Model by the method of Newton Raphson Algorithm. Further, to study its effects for power flow control and losses minimization in the power system. In this paper, optimal placement of UPFC is conducted based on active power loss Sensitivity factors. Based on these sensitivity factors the UPFC is optimally

placed in the required transmission line to investigate the impact of UPFC in the system. The

changes in the system are studied to see the impact of the UPFC. The impact of UPFC are analyzed by using 5-Bus, IEEE 14-bus, and IEEE 30-bus Test systems. The analysis is achieved through developing of software program using MATLAB.

Keywords— LFA (load flow analysis), UPFC (unified power flow control),Z-bus Allocation,transmission loss, Sensitivity analysis.

1. INTRODUCTION

1. UNIFIED POWER FLOW CONTROLLER (UPFC)

Fast progress of power electronics has made Flexible AC Transmission Systems (FACTS) as a promising concept. Researches on FACTS technologies are being performed very actively. Along with advanced control techniques on FACTS devices, power flow among transmission networks is more and more controllable. Among a variety of FACTS controllers, the Unified Power Controller(UPFC) is a new device in FACTS family, which has been introduced by Gyugiy(1991)[13]. It can be used in power systems for several purposes, such as shunt compensation, series compensation, phase shifting, power flow control and voltage control. With the adoption of UPFCs in power systems, the traditional power flow analysis will face new challenges in modeling and solution techniques.

The UPFC consists of two voltage source converters, which are connected back to back through a DC link. The series voltage converter is connected to the transmission line by means of a series transformer and the shunt voltage converter by means of shunt transformer. The series voltage converter injects an AC voltage into the transmission line with controllable magnitude and phase angle. The shunt converter can exchange active and reactive powers with the system, which enables the system to do shunt compensation independently. To investigate the effect of the UPFC on the steady state condition of the system and load flow, different models have been introduced. These models are usually based on modification of traditional load flow methods. The UPFC injection model is easily incorporated in Newton Raphson power flow model to study its effect for power flow control and losses minimization in the power system. The program is written using MATLAB software.

1.1 Basic circuit arrangement of UPFC:

Basically, the unified power flow controller (UPFC) consists of two switching converters. These converters are operated from a common DC link provided by a DC storage capacitor as shown in the Figure 2.

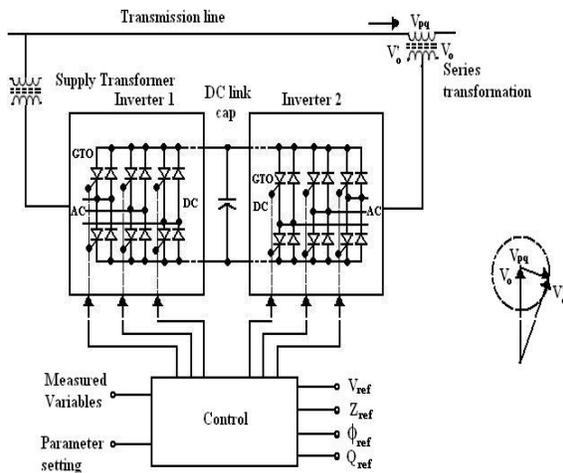


Fig.2. Basic circuit arrangement of UPFC

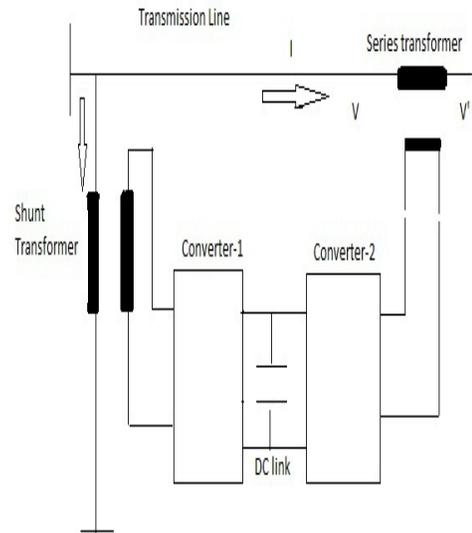
Converter 2 provides the main function of the UPFC by injecting an AC voltage with controllable magnitude and phase angle in series with the transmission line via a series transformer. The basic function of converter 1 is to supply or absorb the real power demand by converter 2 at the common DC link. It can also generate or absorb controllable reactive power and provide independent shunt

reactive compensation for the line. Converter 2 supplies or absorbs locally the required reactive power and exchanges the active power as a result of the series injection voltage.

1.2 UPFC MODEL:

The schematic representation of the UPFC is shown in Fig.3. It consists of two voltage source converters and a dc circuit represented by the capacitor.

Converter 1 is primarily used to provide the real power demand of converter 2 at the common dc link terminal from the ac power system. Converter 1 can also generate or absorb reactive power at its ac terminal, which is independent of the active power transfer to (or from) the dc terminal. Therefore, with proper control, it can also fulfil the function of an independent advanced static VAR compensator providing reactive power compensation for the transmission line and thus executing indirect voltage regulation at the input terminal of the



UPFC.

Fig.3. Schematic diagram of UPFC.

Converter 2 is used to generate a voltage source at the fundamental frequency with variable amplitude ($0 \leq V_T \leq V_{TMAX}$) and phase angle ($0 \leq \phi_T \leq 2\pi$), which is added to the ac transmission line by the series-connected boosting transformer. The inverter output voltage injected in series with line can be used for direct voltage control, series compensation, phase shifter, and their combinations. This voltage source can internally generate or absorb all the reactive power required by

the different type of controls applied and transfers active power at its dc terminal.

1.3 UPFC injection model for power flow studies.

In this study, a model for UPFC, which will be referred as UPFC injection model [1] is derived. This model is helpful in understanding the impact of the UPFC on the power system in the steady state. Furthermore, the UPFC injection model can easily be incorporated in the steady state power flow model. Since the series voltage source converter does the main function of the UPFC, first derive the modelling of a series voltage source converter.

Series connected voltage source converter model: Suppose a series connected voltage source is located between nodes i and j in a power system. The series voltage source converter can be modeled with an ideal series voltage Vs in series with a reactance Xs as shown in fig below,

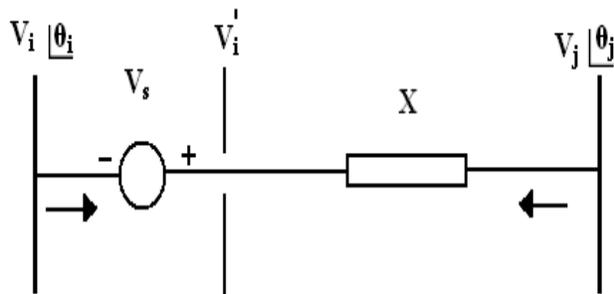


Fig.4. Representation of a series connected VSC

$$V_i^1 = V_s + V_i \tag{1}$$

where V_i^1 = fictitious voltage behind the series reactance.

V_s = series source voltage.

V_i = voltage at i'th node.

The series voltage source V_s is controllable in magnitude and phase, i.e,

$$V_s = rV_i e^{j\gamma} \tag{2}$$

where r = series voltage source coefficient. ($0 < r < r_{max}$)

γ = series voltage source angle. ($0 < \gamma < 2\pi$)

The injection model is obtained by replacing the equivalent circuit of series connected voltage source as Norton's equivalent circuit as shown in fig.5 .The current source,

$$I_s = -jb_s V_s \tag{3}$$

where

$$b_s = 1 / X_s$$

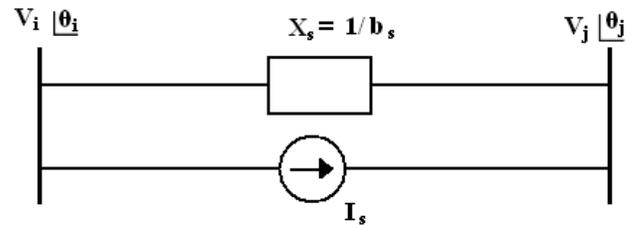


Fig.5. Equivalent Norton's circuit of a series connected VSC.

The power injected into the ith bus

$$\overline{S}_{is} = \overline{V}_i (-\overline{I}_s)^*$$

$$S_{is} = V_i [jb_s r \overline{V}_i e^{j\gamma}]^*$$

$$S_{is} = -b_s r V_i^2 \sin(\gamma) - jb_s r V_i^2 \cos(\gamma)$$

(4)

The power injected into the jth bus

$$\overline{S}_{js} = \overline{V}_j (-\overline{I}_s)^*$$

$$S_{js} = V_j [-jb_s r \overline{V}_i e^{j\gamma}]^*$$

$$S_{js} = b_s r V_i V_j \sin(\theta_{ij} + \gamma) + jb_s r V_i V_j \cos(\theta_{ij} + \gamma)$$

(5)

Where $\theta_{ij} = \theta_i - \theta_j$

From above equations, the injection model of series connected voltage source can be sent as two dependent loads as shown in fig.6.

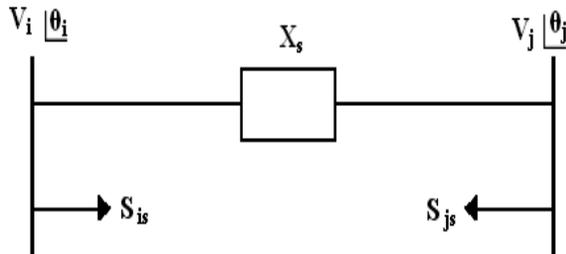


Fig.6 .Injection model for a series connected VSC.

Shunt connected voltage source converter model:

In UPFC, the shunt connected voltage source (converter1) is used mainly to provide the active power, which is injected to the network via the series connected voltage source. When the losses are neglected

$$P_{conv1} = P_{conv2}$$

The apparent power supplied by the series voltage source converter is

$$S_{conv2} = \overline{V_s} \overline{I_{ij}}^* = re^{j\gamma} \overline{V_i} \left[\frac{\overline{V_i} - \overline{V_j}}{jX_s} \right]^* \tag{6}$$

After simplification, the active and reactive power supplied by converter 2 is

$$P_{conv2} = rb_s V_i V_j \sin(\theta_i - \theta_j + \gamma) - rb_s V_i^2 \sin(\gamma)$$

$$Q_{conv2} = -rb_s V_i V_j \cos(\theta_i - \theta_j + \gamma) + rb_s V_i^2 \cos(\gamma) + r^2 b_s V_i^2 \tag{7}$$

The reactive power delivered or absorbed by converter 1 is independently controllable by UPFC and can be modelled as a separate controllable shunt reactive source. In view of above, it is assumed that

$Q_{conv1} = 0$. The UPFC injection model is constructed from the series connected voltage source model with the addition of a power equivalent to $P_{conv1} + j0$ to node i.

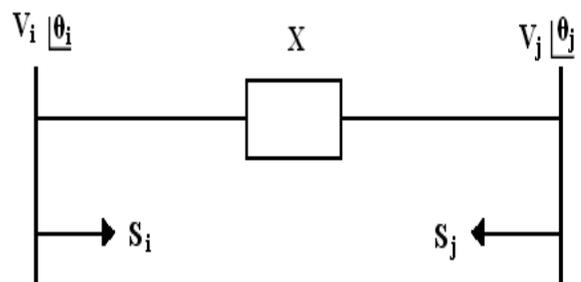


Fig.7 . Complete UPFC model.

Thus, the complete UPFC injection model is shown in fig.7.

$$Q_{si} = rb_s V_i^2 \cos(\gamma) + Q_{shunt}$$

$$Q_{sj} = -rb_s V_i V_j \cos(\theta_{ij} + \gamma)$$

$$P_{si} = rb_s V_i V_j \sin(\theta_{ij} + \gamma)$$

$$P_{sj} = -rb_s V_i V_j \sin(\theta_{ij} + \gamma) \tag{8}$$

1.4 Modification of Jacobian matrix:

The UPFC injection model can easily be incorporated in a load flow program. If a UPFC is located between node i and node j in a power system, the Jacobian matrix is modified by addition of appropriate injection powers. The linearized load flow model is

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} H \dots\dots N \\ J \dots\dots L \end{bmatrix} \begin{bmatrix} \Delta \theta \\ \Delta V / V \end{bmatrix} \tag{9}$$

The Jacobian matrix which is modified is given below, the superscript '0' denotes the Jacobian elements without UPFC.

2. METHODOLOGY

The main objective of the project is to study the effects of Unified Power Flow Controller on loss minimization and loss allocations. To achieve the main objective of the project the UPFC injection model is incorporated in the Newton Raphson algorithm. Thus, in this chapter a detail explanation about UPFC incorporation in load flow is shown. Optimal placement of the UPFC is achieved from active power loss sensitivity factors with respect to the UPFC control parameters. A brief explanation is also given about sensitivity analysis. From the solved load flow model, losses are obtained and these losses are allocated at each bus by the Z-Bus loss allocation method, which is also explained in detail further.

2.1 Incorporation of UPFC in Newton Raphson Power Flow Algorithm:

2.1.1 Newton Raphson Power Flow Algorithm:

From the mathematical modelling point of view, the set of nonlinear, algebraic equations that describe the electrical power network under the steady state conditions are solved for the power flow solutions. Over the years, several approaches have been put forward to solve for the power flow equations. Early approaches were based on the loop equations and methods using Gauss-type solutions. This method was laborious because the network loops has to be specified by hand by the systems engineer. The drawback of these algorithms is that they exhibit poor convergence characteristics when applied to the solution of the networks. To overcome such limitations, the Newton-Raphson method and derived formulations were developed in the early 1970s and since then it became firmly established throughout the power system industry. In the project a Newton Raphson power flow algorithm [2] is used to solve for the power flow problem in a transmission line with UPFC [1]

- **Steps to Incorporate UPFC in Newton-Raphson Algorithm:**

- Step 1: Read the system input data; line data, bus data, generator and load data.

- Step 2: Formation of admittance matrix 'Y' bus of the transmission line between the bus i and j.
- Step 3: Combining the UPFC power equations with network equation, the conventional power flow equation is given as:

$$P_i + jQ_i = \sum_{j=1}^n V_i V_j Y_{ij} \angle(\theta_{ij} - \delta_i + \delta_j) + P_i^1 + jQ_i^1$$

- Step 4: The conventional jacobian matrix are formed due to the inclusion of UPFC. The inclusion of these variables increases the dimensions of the jacobian matrix.
- Step 5: In this step, the jacobian matrix is modified and power equations are mismatched.
- Step 6: The Bus bar voltages are updated at each iteration and convergence is checked. If convergence is not achieved in the next step the algorithm goes back to the step 5 and the jacobian matrix is modified and the power equations are mismatched until convergence is attained.
- Step 7: If the convergence achieved in Step 6, the output load flow is calculated for PQ bus that includes the Bus bar voltages, generation, transmission line flow and losses.

Sensitivity analysis of total active power loss:

- A method based on the sensitivity of the total system active power loss with respect to the control variables of the FACTS device i.e, UPFC is considered.
- For UPFC placed between buses i and bus j, the considered control parameter is the injected series voltage, of controllable magnitude and its phase angle. The active power loss sensitivity factor with respect to these control variables may be given as, loss sensitivity with respect to control parameter of UPFC placed between buses i and bus j.

$$a_{ij} = \frac{\partial P_L}{\partial V_{ij}}$$

- And this can be deduced from the above equation as,

$$\frac{\partial P_L}{\partial V_{ij}} = 2V_i V_j \cos(\delta_i - \delta_j) + 2V_i V_j \sin(\delta_i - \delta_j)$$

- Thus from the above equation the sensitivity factors with respect to active power loss are obtained.

Z-Bus loss allocation method:

- The goal of the Z-bus loss allocation method, is to take a solved power flow and systematically distribute the system transmission losses, among the network buses according to,

$$P_{loss} = \sum_{k=1}^n L_k$$

- The loss component, L_k is the fraction of the system losses allocated to the net real power injection at bus K.
- This is assigned to each individual bus- K, the responsibility of paying for L_k at the market marginal price, λ the extra cost due to loss allocation must then be subtracted from the revenue of the generators and added to the load payments so that the pool remains revenue-neutral.

Z-BUS LOSS ALLOCATION

ALGORITHM:

Step 1: Solve load flow; get bus voltage vector V and total power loss.

Step 2: Obtain bus current vector I from V and complex power injection.

$$S = (P_i + jQ_i)$$

Step 3: Obtain the vector RI.

$$RI = \text{Re} \{Z\} I$$

$$RI = \text{Re} \{Z [\text{Re} (I)]\} + \text{Re} \{Z [\text{Im} (I)]\}$$

Step 4: Calculate the component of total loss due to current injection I_k

at the j^{th} bus.

$$L_k = \Re \left[I_k^* \left(\sum_{j=1}^{nb} R_{kj} \cdot I_j \right) \right]$$

Step 5: Compute L_k for all buses. L_k is the loss allocated to bus K

calculated using the Z-Bus loss allocation method.

Simulation results:

[5-Bus system]:

- The simulation study is done initially for without UPFC device in the 5-Bus test system.

Voltage profile without UPFC.

| Bus numbers | Voltage in p.u | Angle in degree |
|-------------|----------------|-----------------|
| 1 | 1.0600 | 0.0000 |
| 2 | 1.0500 | -2.8470 |
| 3 | 1.0262 | -5.0177 |
| 4 | 1.0257 | -5.3506 |
| 5 | 1.0204 | -6.1727 |

Bus power injections without UPFC.

| Bus numbers | Real power (MW) | Reactive power(MVAR) |
|-------------|-----------------|----------------------|
| 1 | 1.2959 | -0.1274 |
| 2 | 0.2000 | 0.2521 |
| 3 | -0.4500 | -0.1500 |
| 4 | -0.4000 | -0.0500 |
| 5 | -0.6000 | -0.1000 |

Total loss in the system before incorporating UPFC. 4.5895 MW/hr

Loss allocation without UPFC.

| Bus numbers | Distribution of active power loss cost in \$ / hr | Distribution of active power loss cost in / hr |
|-------------------|---|--|
| 1 | 136 | 6800 |
| 2 | 5 | 250 |
| 3 | 19 | 950 |
| 4 | 21 | 1050 |
| 5 | 46 | 2300 |
| Total cost | 229 | 11,450 |

Sensitivity Factors. [5-BUS SYSTEM]

| Line number | From bus | To bus | Sensitivity Factor |
|-------------|----------|----------|--------------------|
| 1 | 1 | 2 | 2.3338 |
| 2 | 1 | 3 | 2.3574 |
| 3 | 2 | 3 | 2.2350 |
| 4 | 2 | 4 | 2.2460 |
| 5 | 2 | 5 | 2.2635 |
| 6 | 3 | 4 | 2.1173 |
| 7 | 4 | 5 | 2.1231 |

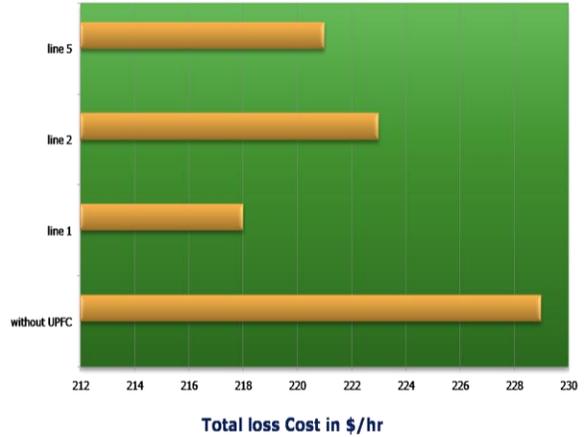
Comparisons of voltage and phase angle profile with UPFC.

| Bus Numbers | UPFC in line-1 | | UPFC in line-2 | | UPFC in line-5 | |
|-------------|----------------|-----------------|----------------|-----------------|----------------|-----------------|
| | Voltage in p.u | Angle in degree | Voltage in p.u | Angle in degree | Voltage in p.u | Angle in degree |
| 1 | 1.0600 | 0.0000 | 1.0600 | 0.0000 | 1.0600 | 0.0000 |
| 2 | 1.0500 | -2.6612 | 1.0500 | -2.8120 | 1.0500 | -2.8495 |
| 3 | 1.0262 | -4.8795 | 1.0264 | -4.9397 | 1.0264 | -4.9926 |
| 4 | 1.0257 | -5.2027 | 1.0259 | -5.2811 | 1.0259 | -5.3175 |
| 5 | 1.0204 | -5.9994 | 1.0205 | -6.1262 | 1.0203 | -6.0282 |

Comparisons of Real & Reactive power injections with UPFC.

5 BUS SYSTEM

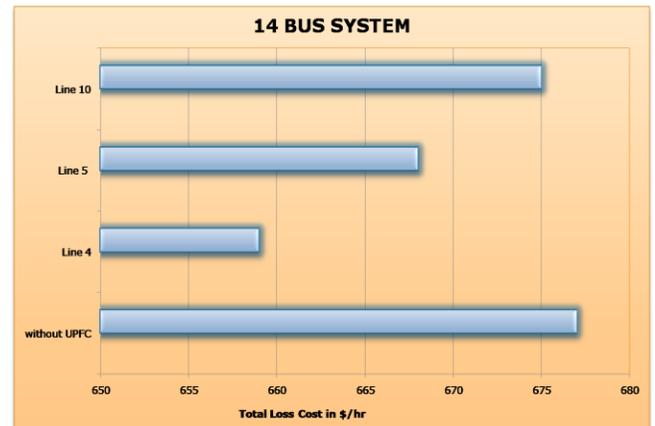
| Bus No | UPFC in line-1 | | UPFC in line-2 | | UPFC in line-5 | |
|--------|-----------------|-----------------------|-----------------|-----------------------|-----------------|-----------------------|
| | Real power (MW) | Reactive power (MVAR) | Real power (MW) | Reactive power (MVAR) | Real power (MW) | Reactive power (MVAR) |
| 1 | 1.2308 | -0.1097 | 1.2796 | -0.1243 | 1.2946 | -0.1280 |
| 2 | 0.2630 | 0.2283 | 0.2000 | 0.2474 | 0.1697 | 0.2504 |
| 3 | -0.4500 | -0.1500 | -0.4346 | -0.1513 | -0.4500 | -0.1500 |
| 4 | -0.4000 | -0.0500 | -0.4000 | -0.0500 | -0.4000 | -0.0500 |
| 5 | -0.6000 | -0.1000 | -0.6000 | -0.1000 | -0.5697 | -0.1017 |



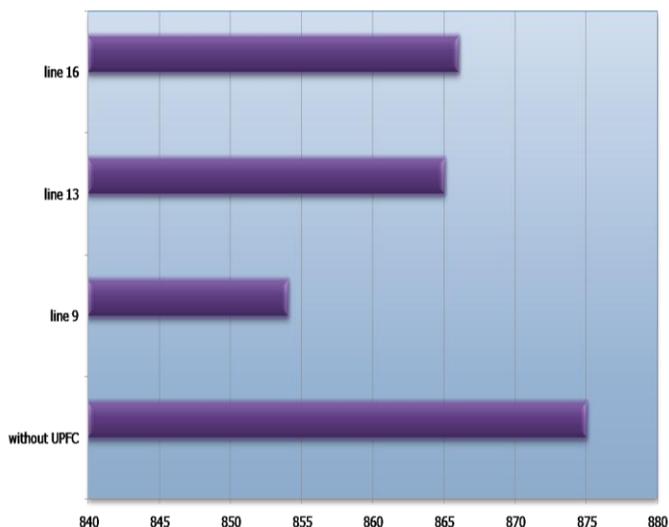
Comparisons of Loss minimization and Allocation with UPFC.

14 BUS SYSTEM

| Bus numbers | UPFC in Line-1 | UPFC in Line-2 | UPFC in Line-5 |
|--------------------------|---|---|---|
| | Distribution of active power loss cost in \$/hr | Distribution of active power loss cost in \$/hr | Distribution of active power loss cost in \$/hr |
| 1 | 125.3190 | 133.2145 | 135.2244 |
| 2 | 7.2450 | 5.1744 | 4.2812 |
| 3 | 19.1835 | 18.1836 | 19.3261 |
| 4 | 20.5882 | 20.4731 | 20.7242 |
| 5 | 45.3440 | 45.9871 | 41.7467 |
| Total cost | 218 | 223 | 221 |
| Total loss in the system | 4.3836 MW | 4.4945 MW | 4.4610 MW |



30 BUS SYSTEM



CONCLUSION:

FACTS devices are conventionally used in the power system for voltage profile improvement, helps real and reactive power flow, enhanced transmission capability. Unified power flow controller is used in this project. In this project, steady state UPFC injection model is incorporated in the load flow model. UPFC is optimally placed using active power loss sensitivity factors which were calculated after performing the load flow analysis. UPFC's role in loss minimization and its influence for loss allocation is verified. Z-Bus loss allocation methodology is used in the project for loss allocation. The impact of UPFC device is tested for 5-bus system, IEEE 14 bus system and IEEE 30 bus system. The results with and without UPFC for the

test systems are tabulated and compared. It is found that generally the system loss would decrease after incorporating UPFC. Thus the loss allocation to each participant would be comparatively lower. MATLAB program was used to cross verify the theoretical results.

Scope For Future Work:

In the project, UPFC's effectiveness for loss minimization is verified. Similarly other FACTS device's role in loss minimization can be checked. In the current project, standard test systems were used. UPFC working in practical systems can be further be evaluated. The loss allocation results using Z-Bus loss allocation can be compared with allocation using alternate methods such as incremental transmission loss allocation, pro-rata and proportional sharing methodologies. A dynamic model of UPFC can be realized for usage in the Optimal Power Flow.

References:

- [1] Loss minimization by incorporation of UPFC in load flow studies, s.v. ravi kumar and s. Siva nagaraju. International journal of electrical and power engineering 1(3):321-327.
- [2] Power Flow Control with UPFC in Power Transmission System, Samina Elyas Mubeen, R. K. Nema, and Gayatri Agnihotri, World Academy of Science, Engineering and Technology 47 2008.
- [3] Load Flow Management in the Interconnected Power Systems Using UPFC Devices C. Bulac Member IEEE, M. Eremia Senior Member IEEE, R. Balaurescu, and V. Ștefănescu.
- [4] Steady State Analysis of Unified Power Flow Controller: Mathematical Modeling and Simulation Studies A.M. Vural, Student Member, IEEE, and M. Tumay.
- [5] Analysis and Modeling of UNIFIED POWER FLOW CONTROLLER: Modification of NEWTON-RAPHSON Algorithm and user-defined modeling approach for POWER FLOW STUDIES,

M. Tumay and A. M. Vural, October 2004 The Arabian Journal for Science and Engineering, Volume 29, Number 2B.

[6] A Novel Method of UPFC Location Based on Sensitivity Factors, Anshul Chaudhary, Manish Jain, P.S.Venkataramu, T.Ananthapadmanabha. Proceedings of the 7th WSEAS International Conference on Power Systems, Beijing, China, September 15-17, 2007.

[7] The modeling of UPFC based on circuit elements in an exact transmission line model, Ali akbar Motiebirjandi Faculty of Electrical Engineering / assistant professor, Kaumars Sabzawari Faculty of Electrical Engineering / M.Sc ,Shahid Rajae University/ Tehran, 67616, Iran.

[8] Study and Effects of UPFC and its Control System for Power Flow Control and Voltage Injection in a Power System, Vibhor Gupta / International Journal of Engineering Science and Technology Vol. 2(7), 2010, 2558-2566.

[9] Control Setting of UNIFIED POWER FLOW CONTROLLER through Load Flow Calculation. C.H Chengaiah, G. V. Marutheswar and R. V. S. Satyanarayana, ARPN Journal of Engineering and Applied Sciences.

[10] "Estimation of UPFC value using Sensitivity Analysis" by Seungwon An, Thomas W. Gedra.

[11] "A unified power flow controller concept for flexible AC transmission systems," in Proc. 5th Int. Conf. AC DC Transmission, London, U.K., Sep. 17–20, 1991, pp. 19–26.

[12] N. G. Hingorani, "Flexible AC transmission," IEEE Spectr., vol. 30, no. 4, pp. 40–45, Apr. 1993.

[13] L.Gyugyi, "A unified power flow control concept for flexible AC transmission systems," Proc. Inst. Elect. Eng., pt. C, vol. 39, no. 4, pp. 323–331, Jul. 1992.

[14] Clodomiro Unsihay and Osvaldo R.Saavedra,"Transmission Loss Unbundling and Allocation Under Pool Electricity Markets"IEEE Trans Power Syst,vol 21,o1,feb 2006.

[15] A.J Conejo et al,"Z Bus Loss Allocation," IEEE Trans. Power Syst., vol 16 no.1,pp.105-110, feb,2001.

[16] F.D Galiana et al,"Incremental Transmission loss allocation under pool dispatch,"IEEE Trans Power Syst,vol 17.no1,pp 26-33,feb 2002.

[17] A.M Leite da silva and J.G.C Costa,"Transmission loss allocation:part 1-single energy market"IEEE Trans Syst.,vol 18,no4,pp 1389-1394,Nov 2003.

[18] J-H Teng,"Power flow and loss allocation for deregulated transmission systems,"Int.J.Elect.Power Energy Syst.,vol 27,no 4,327-333,may 2005.

[19] Optimal location of unified power flow controller (UPFC) in Nigerian Grid System, Mark Ndubuka Nwohu, International journal of electrical and Power Engineering, medwell journals,2010.