Gandhinagar Institute of Technology

We Always Believe in Quality Education

About Gandhinagar Institute of Technology

y was established by Platinum Foundation Trust in December 2006. The Institute is affiliated to Gujarat Technolog
The Trust is registered under Public Trust Act, having Registration No. E-17490 dated 19 June 2006 at Ahmedabac

center for research and engineering study, pursuing knowledge in both fundamental and applied area, and coll
innovation and economic development. The members of the Trust are involved in social activities and are also co
promoting the technical interest of state and country. Gandhinagar Institute of Technology is situated near Village

l, Electronics & Communication, Computer Engineering & Information Technology and Civil Engineering. It also off
ource Management and Information Systems. Gandhinagar Institute of Technology will nurture the intellectual growt
dissemination of knowledge relevant to technology and become one of the premier Engineering and Managemer

Gandhinagar Institute of Technology

We Always Believe in Quality Education

GIT Trustee's

Harishbhai B. Rohera

ifications : B. Com

tion : Member - Board of Trustee & Advisory Board

ground

oprietor: Mahadev Steel Suppliers
o: Vinayak Steel Syndicate
o: Dhiraj Steel Supplier
o: Krishna Steel Trader
vner: National Steel Processor
ustee: Sai Vasant Ghot Darbar
ustee: Jai Jhulelal Mandir

Gahanshyambhai V. Thakkar

ifications : M.A

tion : Member - Board of Trustee & Advisory Board

ground

ofessor at Vivekanand College of Arts, Ahmedabad
M.L.A., Gujarat Assembaly from Mandal
ustee of V.M. Thakkar Charitable Trust, Ahmedabad
nages Muktajeevan Vidhyalaya and BVD High School, Isanpur and Maninagar
visor/Member Kankaria Maninagar Nagarik Sahakari Bank
rector - Adarsh Co-Operative Departmental Stores

Dipakbhai N. Ravani

ifications :B.Com., LL.B.

tion : Member - Board of Trustee & Advisory Board

ground

siness

Pravinbhai A. Shah

ifications : B.A., LL.B.

tion : Member - Board of Trustee & Advisory Board

ground

ident of Zalavad Samaj Jain Seva Trust
rustee of Vasant Atma Charitable Trust
rustee of Rampura Champa Vijya Hospital
rustee of Shantivan & Ambawadi Jain Sangh
rustee of Rampura Kelavani Mandal
rustee of Pampura Panjrapole Trust

Varshaben M. Pandhi

ifications : B.Com

tion : Member - Board of Trustee & Advisory Board

ground

Working experience in the field of Insurance and Investment Advisory for about 20 years

---

EDITORIAL BOARD

Research Methods

Research is another word for gathering of information. The more information we have the closer we get of making our own decision. Research is the result of advancing knowledge created in the past. There are people from all walks of life that contribute to gathered information. These are ordinary people and extraordinary people. They include teachers, students, scientists, professors, scholars, business owners, librarians, book keepers, writers, politicians and many more unknown out there.

Research is designed to solve a particular existing problem so there is a much larger audience eager to support research that is likely to be profitable or solve problems of immediate concern. We also must understand how research impacts our decision making. Most people make decisions without gathered information to back them up. Only few do. Research requires time, effort, and sometimes money to have the evidence you need to make a sound decision that's why many avoid it. The research you do and evidence you gathered will have impact on your future. Be advised, considered the risks or consequences of making an important decision with inadequate evidence.

In conclusion research is very vital to our everyday decision making. It arms you from wrong information and save time and money. It is important to your success as you take on life's challenges and career decisions making. But be careful though, because too much research without action on what you're learning is not good either. The question is how much information is enough? How much information can you afford? Research plus action will most likely guarantee a successful research.

There are five fundamental research methods viz. (1) Experimental methods (2) Correlations (3) Naturalistic observation (4) Survey and (5) Case Study.

Experimental Methods: This method is one in which a researcher manipulates a variable (anything that can vary) under highly controlled conditions to see if this produces (causes) any changes in a second variable. The variable, or variables, that the researcher manipulates is called the independent variable while the second variable, the one measured for changes, is called the dependent variable. Independent variables are sometimes referred to as antecedent (preceding) conditions. All scientific disciplines use this method because they are interested in understanding the laws (cause-and-effect relationships) of nature. The power of the experimental method derives from the fact that it allows researchers to detect cause-and-effect relationships.

In order to see cause-and-effect relationships the researcher must be sure that his manipulations (the independent variable) are the only variables having an effect on the dependent variable. He does this by holding all other variables, variables that might also affect the dependent variable, constant (equivalent, the same). Only by this highly controlled procedure can the researcher be sure that the observed changes in the dependent variable were in fact caused by his manipulations. Experimental studies, therefore, are used when the researcher is interested in determining cause-and-effect relationships. Also, this method can be used when it is appropriate, both practically and ethically, to manipulate the variables.

However, a major limitation is that this method can only be used when it is practical and ethical for the researcher to manipulate the antecedent conditions. A second limitation to this method is that experimental studies are usually done in the highly controlled setting of the laboratory. These conditions are artificial and may not reflect what really happens in the less controlled and infinitely more complex real world.

Correlations: Correlation is classified as a non-experimental, descriptive method. The reason for that is because variables are not directly manipulated as they are in the experimental method. Although correlation is often described as a method of research in its own right, it is really more of a mathematical technique for summarizing data, it is a statistical tool. A co-relational study is one designed to determine the degree and direction of relationship between two or more variables or measures of behavior.

The strength of this method lies in the fact that it can be used to determine if there is a relationship between two variables without having to directly manipulate those variables. In other words, correlation can be used when the experimental method cannot; correlation can be used when it is impractical and/or unethical to manipulate the variables. Correlation also can be used as a basis for prediction.

The greatest limitation of correlation is that it does not tell researchers whether or not the relationship is causal. In other words, correlation does not prove causation. It only shows that two variables are related in a systematic way, but it does not prove nor disprove that the relationship is a cause-and-effect relationship. Only the experimental method can do that.

Naturalistic observation: The naturalistic observation is a type of study classified under the broader category of field studies; non-experimental approaches used in the field or in real-life settings. In the naturalistic observation method the researcher very carefully observes and records some behavior or phenomenon, sometimes over a prolonged period, in its natural setting. The subjects or phenomena are not directly interfered with in any way. In the social sciences this usually involves observing humans or animals as they go about their activities in real life settings. In the natural sciences this may involve observing an animal or groups of animals or some physical phenomena, such as the eruption of a volcano.

The major strength of this method is that it allows researchers to observe behavior in the setting in which it normally occurs rather than the artificial and limited setting of the laboratory. Further uses might include studying nature for its own sake or using nature to validate some laboratory finding or theoretical concept.

One of the limitations is that this is a descriptive method, not an explanatory one. That is, without the controlled conditions of the laboratory, conclusions about cause-and-effect relationships cannot be drawn. Behavior can only be described, not explained. This method can also take a great amount of time. Researchers may have to wait for some time to observe the behavior or phenomenon of interest. Further limitations include the difficulty of observing behavior without disrupting it and the difficulty of coding results in a manner appropriate for statistical analysis.

Survey: The survey, another type of non experimental, descriptive study, does not involve direct observation by a researcher. Rather, inferences about behavior are made from data collected via interviews or questionnaires. Interviews or questionnaires commonly include an assortment of forced-choice questions (e.g. True-False) or open-ended questions (e.g. short answer essay) to which subjects are asked to respond. This sort of data collection is sometimes referred to as a self-report. Surveys are particularly useful when researchers are interested in collecting data on aspects of behavior that are difficult to observe directly and when it is desirable to sample a large number of subjects. Surveys are used extensively in the social and natural sciences to assess attitudes and opinions on a variety of subjects.

The major limitation of the survey method is that it relies on a self-report method of data collection. Intentional deception, poor memory, or misunderstanding of the question can all contribute to inaccuracies in the data. Furthermore, this method is descriptive, not explanatory, and, therefore, cannot offer any insights into cause-and-effect relationships.

Case study: This method is also a non-experimental, descriptive type of study. It involves an in-depth descriptive record, kept by an outside observer, of an individual or group of individuals. This often involves collecting and examining various observations and records of an individual's experiences and/or behaviors. Typical data collected might include biographical data, medical records, family history, observations, interviews, and the results of various psychological tests.

Case studies are particularly useful when researchers want to get a detailed contextual view of an individual's life or of a particular phenomenon. Case studies are also useful when researchers cannot, for practical or ethical reasons, do experimental studies.

This is a descriptive method, not an explanatory one. That is, without the controlled conditions of the laboratory, conclusions about cause-and-effect relationships cannot be drawn. Behavior can only be described, not explained. Case studies also involve only a single individual or just a few and therefore may not be representative of the general group or population. Also, much of the information collected is retrospective data, recollections of past events, and is therefore subject to the problems inherent to memory.

Dr N M Bhatt
Director

Gandhinagar Institute of Technology

We Always Believe in Quality Education

Computer Engineering/Information Technology Engineering Papers

Sr.	Title of Papers	Author	Institution	E-mail ID
1	Graphics Processing Unit With Graphics API	Prof. Darshana Mistry	Assi Prof, Computer Engineering Department, Gandhinagar Institute Of Technology Gandhinagar, India	darhshana.mistry@git.org.in,
2	Data Centic Routing in Wireless Sensor Networks	Mrs.Archana Singh	Lecturer, Computer Engineering Department, Gandhinagar Institute of Technology, Moti-Bhoyani, Kalol	archana.singh@git.org.in
3	Mobile IP (Internet Protocol)	Miss.Maitri Patel, Miss.Mansi Patel, Miss.Happy Patel	Lecturer, Sabar Institute of Technology for Girls, Tajpur	maitri_148@yahoo.co.in mansipatel1488@yahoo.co.in happy_pn88@yahoo.co.in
4	VPN based on Traditional and MPLS	Prof. Rahul A Vaghela	Assi Prof, Information Technology Engineering Department, Gandhinagar Institute of Technology Gandhinagar, India	,rahul.vaghela@git.org.in
5	Performance Comparison of AODV/DSR On-demand Routing Protocols for Mobile Ad Hoc Networks (April – 2011)	Mr. Margil P. Shah Ms. Bhumi S. Jani	Lecturer, Computer Engineering Department, Gandhinagar Institute of Technology, Moti-Bhoyani, Kalol Lecturer, Information Technology Department, Kalol Institute of Technology, Kalol	margil.shah@git.org.in bhumi_mina@yahoo.co.in
6	Semantic Web Mining: Using Ontology Learning and Grammatical Rule Inference Technique	Mr. C. S .Bhatia	Lecturer, Department of Computer Engineering SITG, Gandhinagar	jeet.bhatia@rediffmail.com
7	Software Selection Criteria and Numerical Technique: NPV (Net Present Value)	Prof. Svapnil M Vakharia	Assi Prof, Information Technology Engineering Department, Gandhinagar Institute of Technology Gandhinagar, India	svapnil.vakharia@git.org.in
8	Introduce a New Phase in System Development Life Cycle (SDLC)	Prof. Svapnil M Vakharia	Assi Prof,, Information Technology Engineering Department, Gandhinagar Institute of Technology Gandhinagar, India	svapnil.vakharia@git.org.in
9	Usability Testing: Effective Testing Method To Improve Quality Of Software Design In Technical Education.	Prof. Birendrasinh.K.Zala	Assi Prof,, Information Technology Engineering Department, Gandhinagar Institute of Technology Gandhinagar, India	birendra.zala@git.org.in
10	Oracle Real Application Clusters	Ms. Nisha A Patel	Lecturer, Information Technology Engineering Department, Gandhinagar Institute of Technology Gandhinagar, India	nisha.patel@git.org.in

Gandhinagar Institute of Technology

We Always Believe in Quality Education

Mechanical Engineering Papers

Sr.	Title of Papers	Author	Institution	E-mail ID
1	Application of composite material of Carbon fibre and Glass fibre with Nylon 6 on industrial product	Prof. A.D. Patel, Prof. B S Patel, Prof R S barot	Prof, Mechanical Engineering Department, BVM Engineering college, Vallabh Vidyanagar, Anand, Gujarat, India.	1astorcan@hotmail.com, 2bharatvimlapatel@yahoo.com 3rakesh_le8@yahoo.com
2	Review of Boiling Heat Transfer Correlations for Refrigerants R 22 and R 134a	Prof Nimesh Gajjar Prof. N M Bhatt	Assi Prof , Mechanical Engineering Department, Gandhinagar Institute of Technology, Gandhinagar Director, Gandhinagar Institute of Technology, Gandhinagar	nimesh.gajjar@git.org.in
3	CFD Analysis of an Axial Flow Pump	Prof Tushar Chourushi	Assi Prof , Mechanical Engineering Department, Gandhinagar Institute of Technology, Gandhinagar	tushar.chourushi@git.org.in
4	Design, Improvement and Thrust Bearing Analysis of Oil Expeller Machine	Mr. Mehul.K.Modh Prof J.R.Mevada	U.V. Patel College of Engineering, Kherva, Mehasana.,Gujarat, India Asst. Prof, U. V. Patel College of Engineering, Kherva, Mehasana.,Gujarat, India	modh_mehul@yahoo.co.in
5	Hydrogen – An Alternative Fuel for 21st Century	Mr. Shaival R Parikh	Lecturer, Mechanical Engg. Department, Gandhinagar Institute of Technology	shaival.parikh@git.org.in
6	Analysis of water flow inside the compartment of continuous Electro-deionization (CEDI) unit using Computational fluid dynamics	Mr.Krunal Patel Mr. Girish R. Desaleb Mr.S.K. Thampyb, Mr.Nitin G. Boroleb Mr. Vikas J. Lakheraa	M.Tech Thermal Engg Student, Institute of Technology, Nirma University, Ahemdabad-382 481, INDIA Design Drawing and Workshop Engineering Department, Central Salt and Marine Chemicals Research Institute, Council of Scientific and Industrial Research (CSIR), G.B. Marg, Bhavnagar-364 002,Gujarat, India	krunal_7455@yahoo.com
7	Palm Oil – An Alternative Fuel for I. C. Engines	Mr.Milan J Pandya,	Lecturer, Mechanical Engg. Department, Gandhinagar Institute of Technology	milan.pandya@git.org.in
8	Modeling High Power CO2 Laser Machining Centre with using Autodesk Inventor.	Prof Dhaval P. Patel,	Assi Prof ,Mechanical Engineering Department, Gandhinagar Institute of Technology, Gandhinagar	dhaval.patel@gti.org.in
9	PROJECT SCHEDULING WITH PERT/CPM TECHNIQUES	Prof Priyank V Parekh	Asst. Prof. Mechanical Engineering Dept., Gandhinagar Institute Of Technology	priyank.parikh@git.org.in
10	Filled PTFEs Success Glory for Wear Parts of Reciprocating Gas Compressors over Metals & Alloys	Prof Kedar Bhojaka, Prof Durgesh Jania, Prof Bhoomi Trivedia	Indus Institute of Technology & Engineering, Ahmedabad-382 115.	

Gandhinagar Institute of Technology

We Always Believe in Quality Education

Other Branch Papers

Sr.	Title of Papers	Author	Institution	E-mail ID
1	4G Wireless Technology	Prof . Ketaki Sutariya, Ms. Kruti Lavingia	Assi Prof, Electronics & Communication Dept., Sabar Institute of Tech. for Girls, Tajpur Lecturer, Information Technology Dept., Sabar Institute of Tech. for Girls, Tajpur	ktk_21_06@yahoo.co.in krutilavingia@gmail.com
2	Comparison of Performance Ratings of EDFA and EYCDFA for 1535 nm to 1580 nm Signal Wavelength Span	Mr.Ronak R. Vashi, Mr.Pravin R. Prajapati		
3	Analysis of Self-describing Gridded Geo-science Data Using High Performance Cluster Computing	Ms. Kruti Lavingia, Mr. Nilay Shah	Lecturer, Information Technology Dept., Sabar Institute of Tech. for Girls, Tajpur	krutilavingia@gmail.com nilay@nilayshah.in
4	Microelectromechanical Systems (MEMS) devices, Sensing , Actuation and Applications	Ms. Neelu Kanabar Ms.Binal Gandhi	Lecturer, Department of Instrumentation & Control, Sabar Institute of Technology for Girls At & Po. Tajpur, India	Neelam_kanabar@yahoo.com Sagi.bini@gmail.com
5	Performance Analysis of Antenna Diversity Techniques: A Way to Improve the Bit Error Rate of WiMAX System	Prof. Nirali A. Kotak Mr. Bhavin S. Sedani	Assi Prof, Department of Electronics & Communication Engineering, Gandhinagar Institute of Technology, Ghandhinagar, Gujarat, India Lecturer, Department of Electronics & Communication Engineering, VVP Engineering College of Technology – Vajdi Virda, Rajkot - Gujarat, India	nirali.kotak@git.org.in bhavin_s_sedani@yahoo.com
6	Effect of Solar flares on Background of X-ray Astronomical Objects	Mr.Nirav Y Pandya	Lecturer, Electrical Engg. Department, Gandhinagar Institute of Technology	**nirav.pandya@git.org.in**
7	Seismic Assessment of RC Frame Building with Masonry Infill Wall	Prof. J J Patel. Prof. Raijiwala D. B. Prof M. K Desai.	Assi Prof, Civil Engineering Department, GIT, Moti Bhoyan, Gandhinagar, Gujarat. Prof, Applied Mechanics Department, SVNIT, Surat, Gujarat.	janak.patel@git.org.in
8	Conservation of Energy by LED Lighting & its Advantages	Mr. Sagar J. Kothari	Lecturer, Electrical Engg. Department, Gandhinagar Institute of Technology	sagar.kothari@git.org.in

"Semantic Web Mining: Using Ontology Learning and Grammatical Rule Inference Technique"

C. S .Bhatia[1]
Department of Computer Engineering
SITG, Gandhinagar
jeet.bhatia@rediffmail.com

I.INTRODUCTION

Nowadays the Web has proved to be as a rich and extraordinary data source of information, where multiple domains can be accessed and mined. Mining Web data is referred as **Web Mining**. Some of the objectives of mining web data include finding relevant information discovering new knowledge from web personalized, web synthesis and learning about individual users. Amongst these the most common use is finding relevant information. We simply specify a set of keywords or query as a request or a reference and we get a list of pages, ranked as per similarity of query. Currently searching web face with one problem that many times outcome is not satisfactory because of irrelevance of the information. Searching the exact information from such a huge repository [6] of unstructured web data is still main area of research interest. One solution to this problem is Semantic Web. The Semantic Web is an extension of current Web in which information is given as well defined meaning, hence enabling computers and people to work with better coordination. By using the existing web semantically new semantically structure can be exploited, and then the results of web mining can be improved, thereby building semantic Web [2]. Semantic Web leads the idea of Ontology learning. Ontology learning, through which the outcome leads from the web unstructured data towards the exactness. Many researchers have suggested variety of methods for ontology learning especially concerned to specific data types. In this paper we suggest the process of Ontology learning for the extraction of semantics through Grammatical Rule Extraction Technique.

Keywords: Web Mining, Ontology Learning, Grammar Rule Inference Technique, Semantic Web.

1. Literature Survey

Data mining is the process of extracting previously unknown information from (usually large quantities of) data (text, audio, video etc.), which can, in the right context, lead to knowledge. When data mining techniques are applied to data on Web, we call it as web-data mining or web mining in short. Technically, Web mining refers to the whole of data mining and related techniques that are used to automatically discover and extract information from web documents and services [3]. The web contains huge collection of unstructured data which makes it extremely difficult to search and retrieve valuable information. The size of Web is growing exponentially and there is no force monitoring or handling it. The Web can be seen as the largest database of different types of data in the world. With this huge and ever-growing amount of database, searching exact information directs the attention towards the data mining research. In accordance with **Madria** et al. (1999), we shall divide web-data mining into three categories [19].

-

[1]Chandrapal Singh Bhatia, Associate Professor, Department of Computer Engineering & Information Technology,
Sabar Institute of Technology, Gandhinagar.

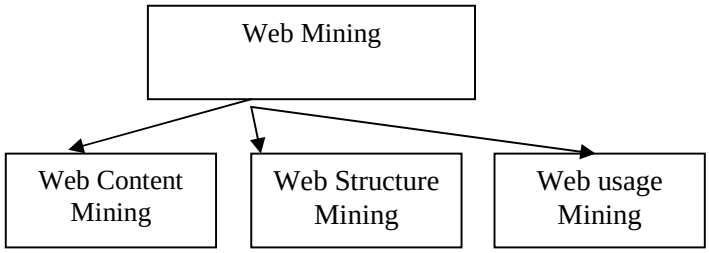


Figure 1: Classification of Web Mining

1. The objective of **Web *content mining*** is to analyze the data contents available in web documents. This includes all types of data text, images, audio file, etc.
2. The objective of **Web *structure mining*** is to focus on link information. It aims to analyze the way in which different web documents are linked together.
3. The objective of **Web *usage mining*** is to analyze the transaction data, which is logged when users interact with the web. Usage mining is sometimes referred to as ***'log mining'***, because it involves mining the web server logs.

In this paper we emphasize on Web Content Mining for text with the objective of achieving exact outcomes with the help of Ontology Learning via Grammatical Rule Extraction Technique. The knowledge provided by ontology is extremely useful in defining the structure and scope for mining Web Content. This paper has been divided in to three parts. First part describes the Semantic Web, the role of Ontologies in Semantic Web and the importance of Semantic Web in ontology Learning. Second part throws light on the approach of Ontology Learning tools and techniques Last part describes how Ontology Learning can be performed using Grammar Rule Extraction Techniques.

II SEMANTIC WEB

When Ontological approach is added in current web then with this extension the current Web is known as, **Semantic Web[4],** It refers to the construction of a machine-understandable semantic layer over the existing web content so as to support better information processing and Web services[1]. The Semantic Web has a developing layered

architecture [9] which is often represented with a diagram first proposed by Tim Berners-Lee[4].
These layers are described as follows:

Unicode and URIs: The concept of Unicode is the standard used for computer character representation. This provide a basis for representing characters used in most of the languages in the world, URIs on the other hand provide the standards for identifying and locating resources (such as pages on the Web).

XML: XML and its related standards, such as Namespaces, and Schemas, form a common medium for structuring data on the Web. They do not communicate the meaning of the data, but nonetheless are well established within the Web [3].

Resource Description Framework: RDF is the first proper layer of the Semantic Web. RDF is a simple metadata representation framework. By using URIs it Identifies Web-based resources, using a graph model that describes relationships between resources.

RDF Schema: a simple type modeling language for describing the classes of resources along with their properties in the basic RDF model. It provides a simple reasoning framework for inferring the types of resources.

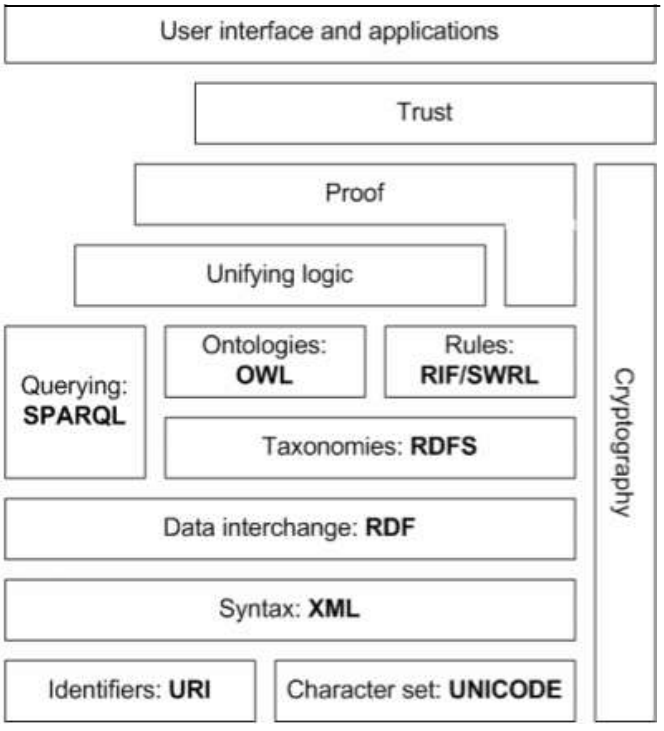


Figure 2: Semantic Web layered architecture

These layers are described as follows:

Unicode and URIs: The concept of Unicode is the standard used for computer character representation. This provide a basis for representing characters used in most of the languages in the world, URIs on the other hand provide the standards for identifying and locating resources (such as pages on the Web).

XML: XML and its related standards, such as Namespaces, and Schemas, form a common medium for structuring data on the Web. They do not communicate the meaning of the data, but nonetheless are well established within the Web already.

Resource Description Framework: RDF is the first proper layer of the Semantic Web. RDF is a simple metadata representation framework. By using URIs it Identifies Web-based resources, using a graph model that describes relationships between resources.

RDF Schema: a simple type modeling language for describing the classes of resources along with their properties in the basic RDF model. It provides a simple reasoning framework for inferring the types of resources.

Ontologies: A richer language for providing more complex constraints on the types of resources and their properties.

Logic and Proof: An (automatic) reasoning system provided on top of the ontology structure to make new inferences. Thus, using such a system, a software agent can make deductions as to whether a particular resource satisfies its requirements or not (and vice versa).

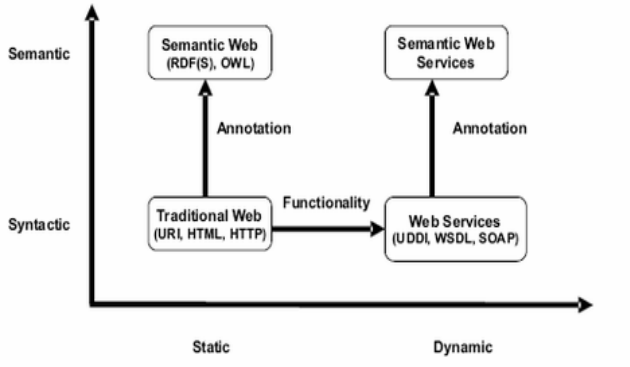


Figure 3: Semantic Web

Trust: The purpose of final layer of the layered architecture is to know trustworthiness of the information by asking questions in Semantic Web. This assures the quality of that information.

The Semantic Web uses in reasoning while searching towards the exactness on web data for the search query hence we can say that Semantic Web helps the Web machine process better. In this paper we are interested in finding out the reasoning from the grammar as used by Semantic Web[12].

III SEMANTIC WEB AND ONTOLOGY

Ontology is "an **explicit specification** of **conceptualization**". [15] In other words, ontology is a domain model which refers *conceptualization* that needs to be *explicitly specified*. According to Borst[201]1997, Ontology is a "formal specification of a shared this minimizes the ambiguities among the concepts. Also, this conceptualization conceptualization". Shared conceptualization means conceptualization should express a *shared* view between several parties, a "*consensus*" rather than an individual view; should be in the format that can easily be understood by machines. In 1998, Studer[17,22] merged these two definitions stating that: *"Ontology is a formal, explicit specification of a shared conceptualization."* [10].In consensual domain models, the primary role of ontologies is to

enhance communication between humans (e.g., establishing a shared vocabulary, explaining the meaning of the shared terms to reach consensus). As formal models, ontologies represent knowledge in a computer process able format thereby enhancing communication between humans and computer programs or two computer programs. Ontologies can be classified according to two dimensions. Foundational (or top-level) ontologies are conceptualizations that contain specifications of domain and problem independent concepts and relations (such as space, time, and matter) based on formal principles derived from linguistics, philosophy, and mathematics. In this thesis we use of the DOLCE foundational ontology. Other examples of top-level ontologies, all discussed and compared in (Borgo et al., 2002), are: the Suggested Upper Merged Ontology (SUMO) (Pease et al., 2002), OpenCyc5 and
the Basic Formal Ontology (BFO) (Smith, 2003).

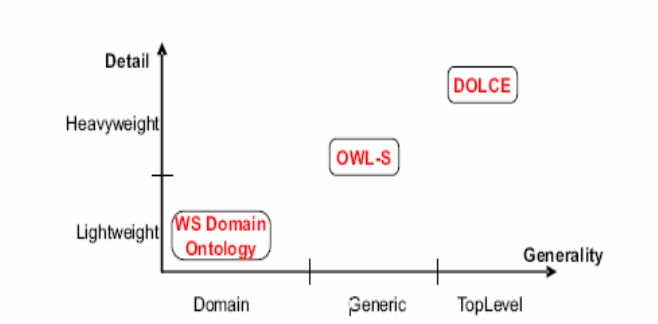


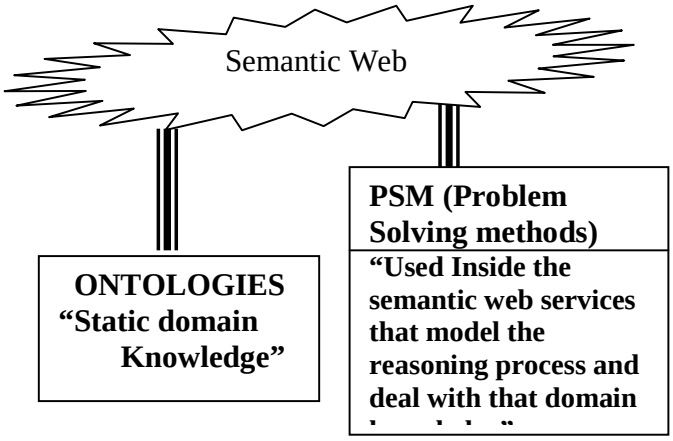
Figure 3 Ontologies classified according to two dimensions

Foundational (or top-level) ontologies are conceptualizations that contain specifications of domain and problem independent concepts and relations (such as space, time, and matter) based on formal principles derived from linguistics, philosophy, and mathematics. In this thesis we use of the DOLCE foundational ontology. Other examples of top-level ontologies, all discussed and compared in (Borgo et al., 2002), are: the Suggested Upper Merged Ontology (SUMO)[23] (Pease et al., 2002), OpenCyc5 and the Basic Formal Ontology (BFO) (Smith, 2003). Generic ontologies contain generic knowledge about a certain domain such as medicine, biology, mathematics or Web services. These domain concepts are often specified in terms of top-level concepts thus inheriting the general theories behind the top-level concepts. In this thesis we used the OWL-S[26] (Martin et al., 2003) ontology which specifies a set of concepts that allow describing Web services. Domain ontologies have the lowest reusability and are specific to a particular domain. Lightweight, domain ontologies are those that are automatically learned; a heavyweight are generic ontology, OWL-S[24], which contains a couple of restrictions but which is still relatively poor from an ontological perspective; a heavyweight, foundational ontology, is DOLCE[11]. In the context of the semantic Web, ontologies improve the exploitation of Web resources by adding a consensual piece of knowledge. Ontology construction is different from the ontology learning as an automatic process. Ontology is an effective conceptualism commonly used for the Semantic Web. There are two main pillars of semantic Web:

(a) Ontologies

(b) PSM

Figure 4 Semantic Webs and Ontology



Ontologies help to output the desired search.

IV. ONTOLOGY LEARNING

Ontology learning (OL) is defined as an approach of building Ontology from knowledge sources using a set of machine learning techniques[18] and knowledge acquisition methods. A unique data source cannot cover all concepts of a target domain of knowledge and since Web is a rich textual source, the Web can be considered as a learning corpus from which domain ontologies are extracted to be used in semantic search systems. Our main objective is to make the semantic search engine more flexible and autonomous to construct domain ontologies from relevant documents in an incremental manner, by combining ontology learning from text and semantic search technology. Ontology learning can be understood in two senses. First one is to develop methods and techniques that allow reducing the effort necessary for the knowledge acquisition process, and second one is in the sense of building ontologies require much time and resources. In this sense we can define Ontology learning as the set of methods and techniques used for building an Ontology from scratch, enriching or adapting an existing ontology in semi automatic fashion using several sources. There is different Ontology learning approaches focus on the type of input used for learning for example ontology learning from text, from dictionary, from knowledge base, from semi structured schemata and from relational schemata. Generally in this paper we are concerned with the ontology learning when input is supplied as text. Whenever text is supplied for the search, the search result depends on how well text has been interpreted by the machine. There may be varieties of text

phrases for similar types of meanings. These are the the ontologies which plays the important role in making the search approachable to the content which matches with the semantics of the phrase. In this way we can say that grammar used in the text phrase plays a important role in the process of the ontology learning. In this paper we hereby point our attention towards the grammar inference rules which can be extracted from the text phrases helping in the process of Ontology Learning[14].

V ONTOLOGY LEARNING USING GRAMMAR INFERENCE TECHNIQUE

The simple process of Ontology learning described in figure shown below the inference of grammar rules are the part of the step of Language Learning. When considering ontologies as general knowledge repositories, ideally reflecting substantial amount of information present on the web, it is obvious that developing them purely manually is infeasible task not only due to the extensive size of data, but also due to the highly dynamic nature of the environment. Therefore the need for automated methods of ontology creation and maintenance is well acknowledged in the community. However, there has been no explicit support for automatically learned ontologies in the main branches of research concerning inference in the Semantic Web. As already stated in the introduction, this thesis concerns the development of a method to learn Ontologies

about document structure, that is, Ontologies that capture the implicit internal structure of a particular document type.

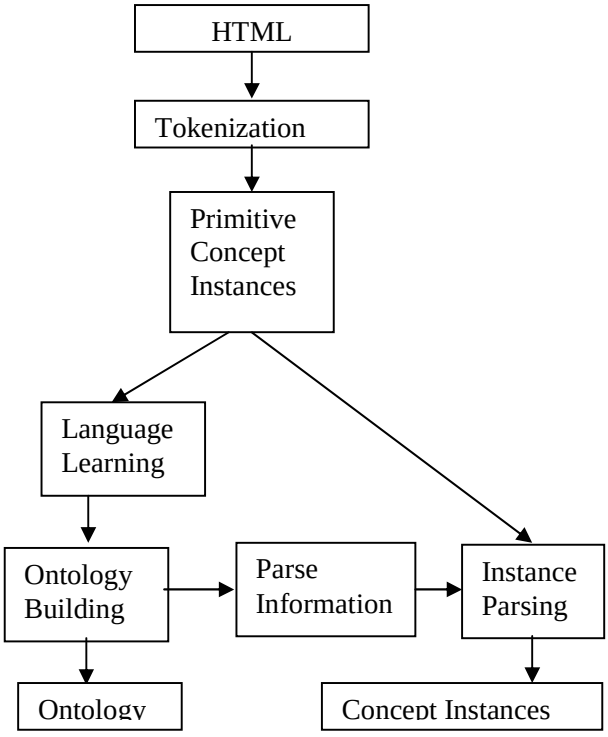


Figure 5 Ontology learning process

As already stated in the introduction, this thesis concerns the development of a method to learn Ontologies about document structure, that is, Ontologies that capture the implicit internal structure of a particular document type. As such, it falls into the category of Ontology learning methods, and most likely in the group of Ontology learning methods from semi-structured data. It can be viewed as learning from semi-structured data because it takes its information not from the actual text of the documents, although some clues come from there, but mainly from the overall structure of the document and in particular from the graphical presentation properties encoded in its HTML code. The extraction of information from web pages is the subject matter of the Wrapper Induction work[13], although their purpose is not the inference of an Ontology, but to enable a database like use of that information. In particular, HTML aware wrapper induction approaches bear similarities to the first step of the method .In particular, they share the preprocessing of HTML pages with the objective of cleaning frequent syntax errors and transformation into valid XML syntax (using the XHTML format) to ease further processing. A particular document type is more than just a sequence of concept instances, requiring some effort to infer its structure. As such, this work applies a regular language learning approach to infer such structure, that is later to be encoded into an Ontology. In particular, it builds on the existing active learning approaches by using the L* algorithm and expands the inference of Deterministic Finite Automata[21] from positive examples, by the application of the Minimum Description Length principle.

Tokenization In the first step we have to map the content of the HTML source documents into a set of primitive concepts. These concepts will be used later as the input for the next step of the ontology learning process. It consists of three steps: (1) Pre-processing, (2) segmentation and (3) classification. The results are stored in a XML file for each source document, where along with each instance text, the concept name and attributes about the instance are also stored.

Pre-processing covers the transformation from HTML to XHTML and fixing, as much as possible, the errors present in the source document. The transformation into XHTML allows us to leverage the existing tools for XML processing. For pre-processing step, the **HTML Tidy 1** [27] tool was used. Documents for which the HTML Tidy was ineffective were not considered in the experiments as they cannot be used in conjunction with XML processing tools and would require either extensive manual corrections or heuristics to recover their content.

Segmentation The segmentation approach is based on the assumption that when a concept instance appears in a document, that is, when a structural change occurs, it will be visible through formatting hints or visual clues. It is further assumed that the concept instances cover, in terms of processable text, the whole document, and that the primitive concept instances do not overlap. In the following descriptions concept instances will be referred to as segments if they are not yet associated with a particular concept name. The source documents, XHTML pages, are trees of tags, with the text placed at the leaves. Many of these tags will be of interest for the extraction of visual clues to segmentation. To organize and select information from these tags and prepare for the classification stage, a set of attributes were chosen to describe

each segment. Segments will therefore be composed of text and a set of attribute value pairs [8].We have used **DOM** [25] API.The set of attributes is divided into two main groups:

1. **XHTML** based attributes - computed from the inspection of XHTML tags. These can further be divided into:

(a) Segment divider - used to drive the segmentation stage.

(b) Non-segment divider.

2**. Post-Processing** - computed after segmentation, either using their segment content, context or classification.In general, the attributes give information about three features of segments:

Text Format These are the attributes that drove the segmentation. In our opinion, these attributes are one of the main means by which people identify the different parts of an article.

Font size (segment divider, post-processing) The current font size. The font size attribute is normalized by dividing it's value by the maximum found in the current article and then converting the result into an integer from 1 to 10.This allows for a better comparison between articles that use a different scale of font sizes on section titles, for example. Font style segment divider) The current font style, that is, if the text is in bold, italic, etc.

Segment divider) the current font style, that is, if the text is in bold, italic, etc.

Text Information A set of attributes that re ect information contained in the segments text. All of them are post-processing attributes.

Text size : The length of the text in each segment has a big weight on its classification. This information is transformed into the following discrete scale: 0 to 30, 30 to 80, 80 to 160, 160 to 1600, 1600 to 16000, bigger than 16000. This scale was obtained empirically. Initially, the first three intervals were collapsed and where intended to capture the size of a one or two line sentence, for example the title of an article. Later it was observed that it was helpful to further sub-divide this interval, so that small segments of text where not mistaken as titles and section titles. The other intervals are there to capture small paragraphs and larger segments of text. Keyword If the text segment starts with one of a given list of keywords, the at-tribute value is the keyword index in the list. Numbering some segments may start with a sequence of numbers, namely the section and subsection titles. This attribute captures the cardinality of that sequence. If a section starts with "1.4.5" the attribute value is 3.

Context Another important factor in the classification of a segment is context information which reacts the segments surrounding the segment being considered. This allows the segment's neighborhood to in influence its classification. Here it was chosen the use of Naïve Bays Classifier, and then we have done classification [5] of given HTML web pages.

Experimental Results

The process was tested on a set of 25 HTML articles from several World Wide Web Conferences[2]. The articles were segmented and them manually classified into the chosen classes (Author(s), Title, AbstractTitle, Abstract- Text, IndexTitle, IndexText, Figure(or Table)Caption, ConferenceTitle, SectionTitle, SubSectionTitle, SubSubSection- Title, SimpleText, FormatedText). These classes represent most of the structure elements present in articles.The keywords used were: "abstract", "figure", "table", "contents" and "overview".Note that the attribute that indicates the classification of the previous segment is updated dynamically during the classification to reflect the result of the actual classification of the previous segment (having a value of "null" for the first segment of each document). During the training phase this attribute is extracted from the previous segment (having a value of "null" for the first segment of each document).The segmentation resulted in 5472 segments which were then used to perform a ten fold cross-validation. The results are presented in tables 1 and 2.

Table 1: Classification Error

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	Avrage
7.68	6.22	2.38	3.47	1.83	2.01	0.73	2.74	0	4.02	3.11

Simple search test

To further validate the results of this method, we developed a simple query language that retrieves documents based on segment contents and their classification. We applied this language to a set of 51 documents, retrieved from the WWWC 2003 conference. This set of documents was segmented and classified using the previously described classifier trained in a separate set of 25 documents[16].

VI .CONCLUSIONS AND FUTURE WORK

In this paper we show how one can perform a robust decomposition of the text of an article into individual segments that correspond to basic natural parts, such as title, section title, text, etc. We aim at automatzing as much as possible the task of building an Ontology of documents. We use a Naive Bays approach to identify re-occurring features in HTML documents.

Once the text of documents is classified according to this ontology we can use this classification to perform sophisticated semantic queries beyond the power of simple we present a simple query language and the preliminary results that show the potential of the approach. As next steps we want to further elaborate the ontology learning process so that we can identify complex concepts, such as a reference, and the relationships between concepts, such as composition relations, and is a relation. An additional direction for future research is the application of the methods described in this paper to documents in other formats, namely PDF and Postscript. In this case, the lack of easily usable tags could be compensated by the use of text positioning information. Now we have seen in Table

1 that fold 9 is error free & fold 10 has 4.02% error and we have calculated average error is 3.11%.In Table 2 we found that maximum error is found in Index Title & Index. Lowest Error is found in Conference Title. With the help of Grammatical Rule Extraction Technique we found that there is

Low percentage of error & whenever we have applied equivalence queries , the results of query is relevant to context,
Which is our objective. With the help of Ontology Learning using Grammatical Inference Technique[7], we have obtained exact results of Query, which is relevant to our text.

CnfT –Conference Title
Ttl – Title
Aths- Authors
ATtl-Abstract Title
ATxt- Abstract Text
Ftxt- Formatted Text
STxt-SimpleText
STtl- SectionTitle
SSTtl – SubSectionTitle
SSSTtl- SubSubSection- Title
IdxTtl- IndexTitle
Idx – IndexText

Fig C- Figure (or Table)Caption

Abbreviation Used in Table 2

Table 2: Classification data – target vs. obtained

	CnfT	Ttl	Aths	ATtl	ATxt	FTxt	STxt	STtl	SSTtl	IdxTtl	Idx	SSSTtl	FigC	Error
CnfT	9	0	0	0	0	0	0	0	0	0	0	0	0	0%
Ttl	0	25	1	0	0	0	0	0	0	0	0	0	0	3.85%
Aths	0	0	18	0	0	5	5	4	0	0	0	0	0	43.75%
ATtl	0	0	0	22	0	1	0	0	0	0	0	0	0	4.35%
ATxt	0	0	0	0	13	0	7	0	0	0	0	0	0	35%
FTxt	0	0	9	2	3	2204	17	3	32	0	0	10	1	3.38%
STxt	0	0	0	0	5	17	2511	0	1	0	0	2	0	0.99%
STtl	0	0	0	0	0	0	0	196	3	0	0	0	0	1.51%
SSTtl	0	0	0	0	0	4	0	9	179	0	0	0	0	6.77%
IdxTtl	0	0	0	0	2	0	0	0	1	0	0	0	0	100%
Idx	0	0	0	0	2	0	1	0	0	0	0	0	0	100%
SSSTtl	0	0	0	0	2	0	0	12	0	0	0	35	1	30%
FigC	0	2	0	0	0	3	1	0	2	0	0	0	88	8.33%

REFERENCES

[1]Aho A. V. & Ulman J.D.: “The theory of Parsing ,Translation and Compiling”,vol.1; Parsing, Prentice-Hal ,Englewood Clins,NJ,1972.

[2]Anglunin D: “Inference of Reversible Languages”, J. ACM 29(1982) 741-765.

[3]Anglunin D.: “On the complexity of minimum inference of regular sets” Inform Control 39 (1978) 337-350.

[4] Berners-Lee, T., Hendler, J. and Lassila, O.: “The Semantic Web,” *Scientific American*. 284(5):35-43, 2001.

[5] Califf M. E. and Mooney R. J: “Relational learning of patter n-match rules for information extraction”. In Proceedings of the Sixteenth National Conference on Artificial Intelligence and Eleventh Conference on Innovative Applications of Artificial Intelligence, Orlando, Florida, USA, pages 328–334, July 1999.

[6] Crescenzi V, Mecca G. and Merialdo P..: “ RoadRunner: Towards automatic data extraction from large web sites”. In VLDB 2001, Proceedings of 27th International Conference on Very Large Data Bases, Roma, Italy, pages 109–118. Morgan Kaufmann, Sept. 1.

[7] DuPont P.:“Regular Grammatical Inference from positive and negative samples by genetic search”: the GIG method, in : Proceedings of second International Colloquium

[8] Embley D. W., Campbell D. M., Jiang Y. S. Liddle S. W. Liddle, Ng, Y.-K., Quass D. and Smith R. D. : “Conceptual-model-based data extraction from multiple-record web pages.” Data and Knowledge Engineering, 31(3):227–251, 1999

[9]Gerber A.J., Barnard A., Van der Merwe A.J. :“Towards a Semantic Web Layered Architecture

[10] Gruber Tom (1993): "A translation approach to portable ontology specifications". In: *Knowledge Acquisition*. 5: 199.

[11]Gold E.M .: “Complexity of automaton identification from given data”, Inform Control 37(1978) 337-350.

[12] Hsu C.-N. and Dung M.-T. : “Generating finite-state transducers for semi-structured data extraction from the web” Information Systems, 23(8):521–538, 1998.

[2]WWWC96, WWWC2000, WWWC2001, WWWC2002 and WWWC2003

[13] Kushmerick N.. : “Wrapper induction: Efficiency and expressiveness.” Artificial Intelligence, 118(1–2):15–68, Apr. 2000.

[14] Laender A. H. F., . Ribeiro-Neto B. A. Ribeiro-Neto, A. S. da Silva, and J. S. Teixeira : “A brief survey of web data extraction tools.” SIGMOD Rec., 31(2):84–93, 2002.

[15] Lars Marius Garshol (2004):*Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all* on www.ontopia.net. 13 October 2008.

[16]Martins A., Sofia H. Pinto and Oliveira A: “Using a more Powerful Teacher to reduce the Number of Queries of the L* Algorithm” .

[17] Maedche A, Stabb S. (2001) : “Ontology Learning for the Semantic Web”. IEEE intelligent Systems, Special Issue on the semantic Web, 16(2)

[18] Mitchell, T. (1997): *Machine Learning*, McGraw Hill. ISBN 0-07-042807-7.

[19] Madria S.,Bhowmick S.Sourav,Ng W.K.& Lim E.M.: “Research Issues in Web Data Mining”,Citeseer[X] Beeta,1999,p.4-12.

[20] Noy, N., McGuinness, D.: “Ontology development 101: A guide to creating your first ontology”. Technical Report SMI-2001-0880, Stanford Medical Informatics (SMI), Department of Medicine, Stanford University School of Medicine (2001).

[21] Perakath C. Benjamin et al. (1994): "*IDEF5 Method Report*." Knowledge Based Systems, Inc.
[22] Studer, R., Benjamins, V., Fensel, D.: Knowledge engineering, principles and methods. Data and Knowledge Engineering **25** (1998) 161–197
[23] Sowa, John F : "Top-level ontological categories." In International Journal of Human-Computer Studies, v. 43 (November/December 1995) p. 669-85.
[24] Laender A. H. F., . Ribeiro-Neto B. A. Ribeiro-Neto, A. S. da Silva, and J. S. Teixeira : "A brief survey of web data extraction tools." SIGMOD Rec., 31(2):84–93, 2002.
[25] W3C. The Document Object Model. http://www.w3.org/DOM
[26] OWL. (http://www.w3.org/TR/2003/ CR-owl-features-20030818/)
[27] http://www.w3.org/People/Raggett/tidy/

DATA CENTRIC ROUTING IN WIRELESS SENSOR NETWORKS

Archana Singh
Lecturer in Computer Engineering Department, Gandhinagar Institute of Technology, Moti-Bhoyani, Kalol

ABSTRACT: The emergence of Wireless Sensor Networks (WSN) as one of the dominant technology trends in the coming decades has posed numerous challenges to researchers. These networks are likely to be comprised of hundreds, and potentially thousands of tiny sensor nodes, functioning autonomously, and in many cases, without access to renewable energy resources.
While the set of challenges in sensor networks are diverse, we focus on fundamental networking challenges (routing) in sensor networks. Routing in WSN is very challenging due to the inherent characteristics that distinguish these networks from other wireless networks like mobile ad hoc networks or cellular networks. In this paper we present two routing algorithm used by WSN, namely Sensor Protocols for Information via Negotiation (SPIN) and Directed Diffusion. Both these algorithms are multi hop flat routing protocols and are data centric. These two protocols motivated the design of many other protocols which follow a similar concept. Advantages and performance issues of these two algorithms are also highlighted.
Keywords: Wireless Sensor Networks, Data Centric Routing, Sensor Protocols for Information via Negotiation, Directed Diffusion.

1. INTRODUCTION

Due to recent technological advances, the manufacturing of small and low cost sensors became technically and economically feasible. A Wireless Sensor Network (WSN) contains hundreds or thousands of these sensor nodes. They greatly extend our ability to monitor and control the physical environment from remote locations. The sensing electronics measure ambient condition related to the environment surrounding the sensor and transforms them into an electric signal. When networked, sensors can aggregate such data to provide a rich, multi-dimensional view of the environment. Each sensor node operates autonomously with no central point of control in the network, and each node bases its decisions on its mission, the information it currently has, and its knowledge of its computing, communication and energy resources.

However, sensor nodes are constrained in energy supply and bandwidth. Thus, innovative techniques that eliminate energy inefficiencies that

would shorten the lifetime of the network are highly required. Such constraints combined with a typical deployment of large number of sensor nodes pose many challenges to the design and management of WSNs and necessitate energy-awareness at all layers of the networking protocol stack. At the network layer, it is highly desirable to find methods for energy-efficient route discovery and relaying of data so that the lifetime of the network is maximized. Routing in WSNs is very challenging due to the inherent characteristics that distinguish these networks from other wireless networks like mobile ad hoc networks or cellular networks.

2. ROUTING CHALLENGES IN WSN's

Despite the innumerable applications of WSNs, these networks have several restrictions, e.g., limited energy supply, limited computing power, and limited bandwidth of the wireless links connecting sensor nodes. One of the main design goals of WSNs is to carry out data communication while trying to prolong the lifetime of the network and prevent connectivity degradation by employing aggressive energy management techniques. The design of routing protocols in WSNs is influenced by many challenging factors.

- *Energy consumption without loosing accuracy:* Because wireless sensors have a limited supply of energy, energy-conserving communication protocols and computation are essential.
- *Computation:* Sensors have limited computing power and therefore may not be able to run sophisticated network protocols.
- *Fault Tolerance*: Some sensor nodes may fail or be blocked due to lack of power, physical damage, or Environmental interference. The

failure of sensor nodes should not affect the overall task of the sensor network.

- *Scalability*: The number of sensor nodes deployed in the sensing area may be in the order of hundreds or thousands, or more. Any routing scheme must be able to work with this huge number of sensor nodes.
- *Communication:* The bandwidth of the wireless links connecting sensor nodes is often limited, on the order of a few hundred Kbps, further constraining inter-sensor communication.
- *Data Aggregation*: Since sensor nodes may generate significant redundant data, similar packets from multiple nodes can be aggregated so that the number of transmissions is reduced. Data aggregation is the combination of data from different sources according to a certain aggregation function, e.g., duplicate suppression, minima, maxima and average.

3. WHAT DOES DATA CENTRIC MEAN

"Data-centric", means routing is based on data contained in the sensor nodes rather than traditional IP theme where end-to-end delivery method is used based on unique identifications. There is no network layer node identification used for routing. Message exchange is driven by the application layer (using link layer support). Use of data naming allows negotiation between nodes concerning what data to forward to eliminate redundancy, and the ability to aggregate primitive data into more complex data.

There are two data centric routing protocols devised for Wireless Sensor Networks. They are:

1) Sensor Protocols for Information via Negotiation (SPIN)

2) Directed Diffusion

Along with being data-centric, both SPIN and directed diffusion are also *flat based routing protocols*. In flat networks, each node typically plays the same role and sensor nodes collaborate together to perform the sensing task. Due to the large number of such nodes, it is not feasible to assign a global identifier to each node. This consideration has led to data centric routing, where the Base Sensor sends queries to certain regions and waits for data from the sensors located in the selected regions. Since data is being requested through queries, attribute-based naming is necessary to specify the properties of data. Early works on data centric routing, e.g., SPIN and directed diffusion were shown to save energy through data negotiation and elimination of redundant data. These two protocols motivated the design of many other protocols which follow a similar concept. In the rest of this subsection, we summarize these protocols and highlight their advantages and their performance issues.

4. SENSOR PROTOCOLS FOR INFORMATION VIA NEGOTIATION (SPIN)

SPIN is a sender initiated routing protocol. It is a family of adaptive protocols that was designed to disseminate individual sensor observations to all sensors in a network, treating all sensors as potential base stations (sink nodes). SPIN, thus, provides a way of replicating complete views of the environment throughout an entire network. This enables a user to query any node and get the required information immediately.

Conventional protocols like flooding or gossiping based routing protocols waste energy and bandwidth when sending extra and un-necessary copies of data by sensors covering overlapping areas. The drawbacks of flooding include *implosion*, which is caused by duplicate messages sent to the same node, *overlap* when two nodes sensing the same region will send similar packets to the same neighbor and *resource blindness* by consuming large amounts of energy without consideration for the energy constraints.

To overcome the problems of implosion and overlap, SPIN nodes negotiate with each other before transmitting data. Negotiation helps ensure that only useful information will be transferred. To negotiate successfully, however, nodes must be able to describe or name the data they observe. We refer to the descriptors used in SPIN negotiations as *meta-data*. The use of meta-data descriptors eliminates the possibility of overlap because it allows nodes to name the portion of the data that they are interested in obtaining. In SPIN, nodes poll their resources before data transmission. Each sensor node has its own *resource manager* that keeps track of resource consumption; applications probe the manager before transmitting or processing data. This allows sensors to cut back on certain activities when energy is low.

4.1 SPIN PHILOSOPHY

SPIN family of protocols rests upon two basic ideas.
1) Exchanging sensor data may be an expensive network operation, but exchanging data about sensor data need not be.
2) Nodes in a network must monitor and adapt to changes in their own energy resources to extend the operating lifetime of the system.

4.2 META-DATA

Sensors use meta-data to succinctly and completely describe the data that they collect. If *x* is the meta-data descriptor for sensor data *X*, then the size of *x* in bytes must be shorter than the size of *X*, for SPIN to be beneficial. If two pieces of actual data are distinguishable, then their corresponding meta-data should be distinguishable. SPIN does not specify a format for meta-data; this format is application-specific.

4.3 SPIN MESSAGES

SPIN nodes use three types of messages to communicate:

- ADV – new data advertisement. When a SPIN node has data to share, it can advertise this fact by transmitting an ADV message containing meta-data.
- REQ – request for data. A SPIN node sends an REQ message when it wishes to receive some actual data.
- DATA – data message. DATA messages contain actual sensor data with a meta-data header.

ADV and REQ messages contain only meta-data and are smaller than their corresponding DATA messages. ADV and REQ messages will therefore be cheaper to transmit and receive than their corresponding DATA messages.

4.4 SPIN PROTOCOLS

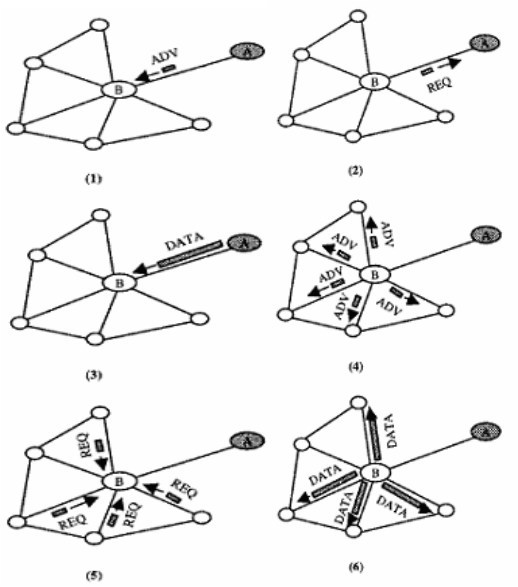
There are mainly four protocols that follow the SPIN philosophy outlined in the previous section.

- *SPIN-PP* – tackle the basic problem of data transmission under ideal conditions, solves problems for net works using point-to-point transmission
- *SPIN-EC*, an energy-conserving version of SPIN-PP
- *SPIN-BC*, tackle the basic problem of data transmission under ideal conditions, where energy is plentiful and packets are never lost, solves this problem for networks using broadcast media.
- *SPIN-RL*, a reliable version of SPIN-BC

4.4.1 SPIN-PP: A THREE-STAGE HANDSHAKE PROTOCOL FOR POINT-TO-POINT MEDIA

The first SPIN protocol, SPIN-PP, is optimized for networks using point-to-point transmission media, where it is possible for nodes *A* and *B* to communicate exclusively with each other without interfering with other nodes The SPIN-PP protocol works in three stages (ADV–REQ–DATA). The protocol starts when a node advertises new data that it is willing to disseminate. It does this by sending an ADV message to its neighbors, naming the new data (ADV stage). Upon receiving an ADV, the neighboring node checks to see whether it has already received or requested the advertised data. If not, it responds by sending an REQ message for the missing data back to the sender (REQ stage). The protocol completes when the initiator of the protocol responds to the REQ with a DATA message, containing the missing data (DATA stage).

Figure shows an example of the protocol. Upon receiving an ADV packet from node A, node B checks to see whether it possesses all of the advertised data (1). If not, node B sends an REQ message back to A, listing all of the data that it would like to acquire (2). When node A receives the REQ packet, it retrieves the requested data and sends it back to node B as a DATA message (3). Node B, in turn, sends ADV messages advertising the new data it received from node A to all of its neighbors (4). It does not send an advertisement back to node A, because it knows that node A already has the data. These nodes then send advertisements of the new data to all of their neighbors, and the protocol continues.



4.4.2 SPIN-EC: SPIN-PP WITH A LOW-ENERGY THRESHOLD

The SPIN-EC protocol adds a simple energy-conservation heuristic to the SPIN-PP protocol. When energy is plentiful, SPIN-EC nodes communicate using the same three-stage protocol as SPIN-PP nodes. When a SPIN-EC node observes that its energy is approaching a low-energy threshold, it adapts by reducing its participation in the protocol. In general, a node will only participate in a stage of the protocol if it believes that it can complete all the other stages of the protocol without going below the low-energy threshold. Similarly, if a node receives an advertisement, it does not send out a request if it does not have enough energy to transmit the request and receive the corresponding data. This approach does not prevent a node from receiving, and therefore expending energy on, ADV or REQ messages below its low-energy threshold. It does, however, prevent the node from ever handling a DATA message below this threshold.

4.4.3 SPIN-BC: A THREE-STAGE HANDSHAKE PROTOCOL FOR BROADCAST MEDIA

In broadcast transmission media, nodes in the network communicate using a single, shared channel. As a result, when a node sends out a message in a lossless, symmetric broadcast network, it is received by every node within a certain range of the sender, regardless of the message's destination. SPIN-BC improves upon SPIN-PP for broadcast networks by exclusively using cheap, one-to-many communication. This means that all messages are sent to the broadcast address and thus processed by all nodes that are within transmission range of the sender.

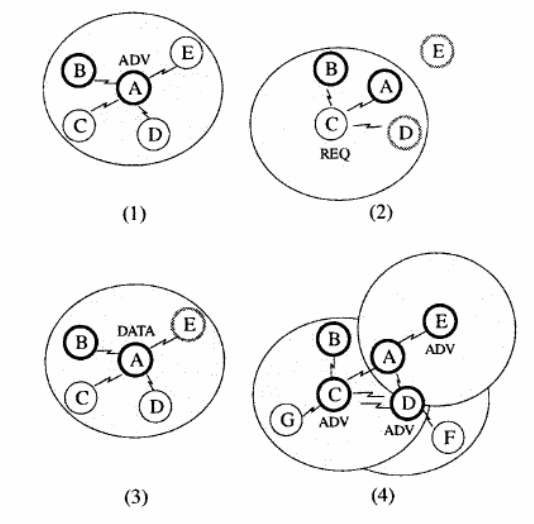
Like the SPIN-PP protocol, the SPIN-BC protocol has an ADV, REQ, and DATA stage, which serve the same purpose as they do in SPIN-PP. There are three central differences between SPIN-PP and SPIN-BC.

1) All SPIN-BC nodes send their messages to the broadcast address, so that all nodes within transmission range will receive the messages.

2) SPIN-BC nodes do not immediately send out requests when they hear advertisements for data they need. Upon receiving an ADV, each node checks to see whether it has already received or requested the advertised data. If not, it sets a random timer to expire, uniformly chosen from a predetermined interval. When the timer expires, the node sends an REQ message out to the broadcast address, specifying the original advertiser in the header of the message. When nodes other than the original advertiser receive the REQ, they cancel their own request timers, and prevent themselves from sending out redundant copies of the same request.

3) SPIN-BC node will send out the requested data to the broadcast address once and only once, as this is sufficient to get the data to all its neighbors. It will not respond to multiple requests for the same piece of data.

Figure 4 shows an example of the protocol. Upon receiving an ADV packet from node A, A's neighbors check to see whether they have received the advertised data (1). Three of A's neighbors, C, D, and E, do not have A's data, and enter request suppression mode for different, random amounts of time. C's timer expires first, and C broadcasts a request for A's data (2), which in turn suppresses the duplicate request from D. Though several nodes receive the request, only A responds, because it is the originator of the ADV packet (3). After A sends out its data, E's request is suppressed, and C, D, and E all send out advertisements for their new data (4).



4.4.4 SPIN-RL: SPIN-BC FOR LOSSY NETWORKS

SPIN-RL, a reliable version of SPIN-BC, can disseminate data efficiently through a broadcast network, even if the network loses packets or communication is asymmetric. The SPIN-RL protocol incorporates two adjustments to SPIN-BC to achieve reliability.

- Each SPIN-RL node keeps track of which advertisements it hears from which nodes, and if it does not receive the data within a reasonable period of time following a request, the node rerequests the data. It fills out the originating-advertiser field in the header of the REQ message with a destination, randomly picked from the list of neighbors that had advertised that specific piece of data.

- SPIN-RL nodes limit the frequency with which they will resend data. If a SPIN-RL node sends out a DATA message corresponding to a specific piece of data, it will wait a predetermined amount of time before responding to any more requests for that piece of data.

4.5 PERFORMANCE COMPARISON OF SPIN PROTOCOLS

Using the simulator NS-2, comparison was performed between SPIN protocols with classic flooding and gossiping and the ideal data distribution protocol.

In terms of time, SPIN-PP achieves comparable results to classic flooding protocols, and in some cases outperforms classic flooding. In terms of energy, SPIN-PP uses only about 25% as much energy as a classic flooding protocol. SPIN-EC is able to distribute 60% more data per unit energy than flooding. In all of our experiments, SPIN-PP and SPIN-EC outperformed gossiping. They also come close to an ideal dissemination protocol in terms of both time and energy under some conditions. SPIN-BC and SPIN-RL are able to use one-to-many communications exclusively, while still acquiring data faster than flooding using less energy. Not only can SPIN-RL converge in the presence of network packet losses, it is able to dissipate twice the amount of data per unit energy as flooding.

4.6 ADVANTAGES:

- Meta-data negotiation and resource adaptation. SPIN provides much energy savings than flooding and meta- data negotiation almost halves the redundant data.
- Maintains only local information about the nearest neighbors.
- Suitable for mobile sensors since the nodes base their forwarding decisions on local neighborhood information

4.7 DISADVANTAGES

- It cannot isolate the nodes that do not want to receive information; unnecessary power may be consumed.
- SPIN's data advertisement mechanism cannot guarantee the delivery of data.

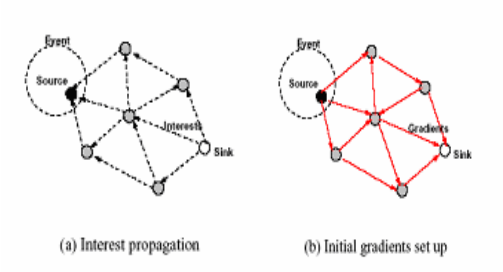
4.7 SUGGESTIONS/IMPROVEMENTS /FUTURE WORK

- Study SPIN protocols in mobile wireless network models
- Develop more sophisticated resource-adaptation protocols to use available energy well
- Design protocols that make adaptive decisions based not only on the cost of communicating data, but also the cost of synthesizing it

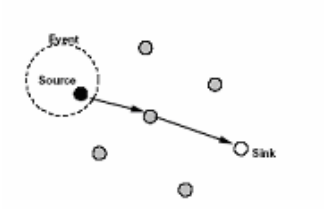
5. DIRECTED DIFFUSION

Directed Diffusion is receiver-initiated routing protocol Directed diffusion consists of several elements: interests, data messages,

gradients, and reinforcements. An *interest* message is a query or an interrogation which specifies what a user wants. Each interest contains a description of a sensing task that is supported by a sensor network for acquiring data. Typically, *data* in sensor networks is the collected or processed information of a physical phenomenon. Such data can be an *event* which is a short description of the sensed phenomenon. In directed diffusion, data is *named* using attribute-value pairs. A sensing task (or a subtask thereof) is disseminated throughout the sensor network as an *interest* for named data. This dissemination sets up *gradients* within the network designed to "draw" events (*i.e., data* matching the interest). Specifically, a *gradient* is direction state created in each node that receives an interest. The gradient direction is set toward the neighboring node from which the interest is received. Events start flowing towards the originators of interests along multiple gradient paths. The sensor network *reinforces* one, or a small number of these paths. Figure 1 illustrates these elements.



(a) Interest propagation (b) Initial gradients set up



(c) Data delivery along reinforced path

5.1 NAMING

In directed diffusion, task descriptions are *named* by, for example, a list of attribute-value pairs that describe a task. Intuitively, the task description specifies an interest for data matching the attributes. For this reason, such a task description is called an *interest*. Given a set of tasks supported by a sensor network, then, selecting a naming scheme is the first step in designing directed diffusion for the network.

5.2 INTERESTS AND GRADIENTS

The named task description of Section constitutes an *interest.* An interest is usually injected into the network at some (possibly arbitrary) node in the network. We use the term *sink* to denote this node.

5.3 INTEREST PROPAGATION

This sink node records the task and for each active task, the sink periodically *broadcasts* an interest message to each of its neighbors. Intuitively, this initial interest may be thought of as *exploratory*; it tries to determine if there indeed are any sensor nodes that detect the specified task. To do this, the initial exploratory interest specifies a low data rate. The interest is soft state and it will be periodically refreshed by the sink.

Every node maintains an interest cache. Each item in the cache corresponds to a *distinct* interest. Two interests are distinct if their attribute differs. . The definition of distinct interests also allows interest *aggregation*. Interest entries in the cache *do not contain information about the sink* but just about the immediately previous hop. An entry in the interest cache has several fields. A timestamp field indicates the timestamp of the last received matching interest. The interest entry also contains several gradient fields, up to one per neighbor. A *gradient* contains

- the node where to forward data (the *previous hop* where the *interest* was received from)
- Data rate, which tells how often data events should be forwarded
- duration, which tells how long data should be forwarded

When a node receives an interest, it checks to see if the interest exists in the cache. If no matching entry exists, the node creates an interest entry. The parameters of the interest entry are instantiated from the received interest. This entry has a single gradient towards the neighbor from which the interest was received, with the specified event data rate. If there exists an interest entry, but no gradient for the sender of the interest,

the node adds a gradient with the specified value. Finally, if there exists both an entry and a gradient, the node simply updates the timestamp and duration fields. When a gradient expires, it is removed from its interest entry.

After receiving an interest, a node may decide to re-send the interest to some subset of its neighbors. To its neighbors, this interest appears to originate from the sending node, although it might have come from a distant sink. This is an example of a local interaction. In this manner, interests diffuse throughout the network. Not all received interests are re-sent. A node may suppress a received interest if it recently re-sent a matching interest.

5.4 GRADIENT ESTABLISHMENT

Every pair of neighboring nodes establishes a gradient towards each other. This is a crucial consequence of local interactions. In summary, interest propagation sets up state in the network (or parts thereof) to facilitate "pulling down" data towards the sink.

5.5 DATA PROPAGATION

A sensor node that is within the specified target region processes interests as described above. In addition, the node tasks its local sensors to begin collecting samples. A sensor node that detects a target searches its interest cache for a matching interest entry. When it finds one, the node tasks its sensor subsystem to generate event. The source then sends to each neighbor for whom it has a gradient. This *data* message is, in effect unicast individually to the relevant neighbors. A node that receives a data message from its neighbors attempts to find a matching interest entry in its cache. If no match exists, the data message is silently dropped. If a match exists, the node checks the *data cache* associated with the matching interest entry. This cache keeps track of recently seen data items. It has several potential uses, one of which is loop prevention. If a received data message has a matching data cache entry, the data message is silently dropped. Otherwise, the received message is added to the data cache and the data message is re-sent to the node's neighbors.

5.6 REINFORCEMENT FOR PATH ESTABLISHMENT AND TRUNCATION

The sink initially and repeatedly diffuses an interest for a low-rate event notification. These are called *exploratory* events, since they are intended for path setup and repair. The gradients set up for exploratory events *exploratory* gradients. Once a source detects a matching target, it sends exploratory events, possibly along multiple paths, towards the sink. After the sink starts receiving these exploratory events, it *reinforces* one particular neighbor in order to "draw down" real *data* (*i.e.*, events at a higher data rate that allow high quality tracking of targets). The gradients set up for receiving high quality tracking events *data* gradients.

5.7 PATH ESTABLISHMENT USING POSITIVE REINFORCEMENT

In general, this novel feature of directed diffusion is achieved by *data driven* local rules. One example of such a rule is to reinforce any neighbor from which a node receives a previously unseen event. To reinforce this neighbor, the sink re-sends the original interest message but with a smaller interval (higher data rate). When the neighboring node receives this interest, it notices that it already has a gradient towards this neighbor. Furthermore, it notices that the sender's interest specifies a higher data rate than before. If this new data rate is also higher than that of any existing gradient (intuitively, if the "outflow" from this node has increased), the node must also reinforce at least one neighbor. Again, the same local rule choices apply. For example, this node might choose that neighbor from whom it first received the latest event matching the interest. Alternatively, it might choose all neighbors from which new events were recently received. The local rule, then, *selects an empirically low-delay path.*

It is very reactive to changes in path quality; whenever one path delivers an event faster than others, the sink attempts to use this path to draw down high quality data. However, because it is triggered by receiving one new event, this could be wasteful of resources. More sophisticated local rules are possible, including choosing that neighbor from whom the most events have been received, or that neighbor which *consistently* sends events before other neighbors. These choices trade off reactivity for increased stability.

5.8 PATH TRUNCATION USING NEGATIVE REINFORCEMENT

The algorithm described above can result in more than one path being reinforced. For

example if the sink reinforces neighbor A, but then receives a new event from neighbor B, it will reinforce the path through B. If the path through B is consistently better (*i.e.,* B sends events before A does), we need a mechanism to *negatively reinforce* the path through A. One mechanism for negative reinforcement is soft state, *i.e.,* to time out all data gradients in the network unless they are explicitly reinforced. With this approach, the sink would periodically reinforce neighbor B, and cease reinforcing neighbor A. All gradients along the path through A would eventually degrade to being exploratory gradients.

5.9 ADVANTAGES

- Data-centric dissemination
- Robust multi-path delivery
- Reinforcement-based adaptation to the empirically best network path
- Energy savings with in-network data aggregation and caching
- Gives designers the freedom to attach different semantics to gradient values
- Reinforcement can be triggered not only by sources but also by intermediate nodes

5.10 DISADVANTAGES

- It may consume memory since all the attribute list is being sent

5.11SUGGESTIONS / IMPROVEMENTS /FUTURE WORK

- Exploration of possible naming schemes

6. COMPARISON BETWEEN THE TWO PROTOCOLS

Both SPIN and Directed Diffusion protocols are optimized for disseminating application specific information in a sensor network. Both SPIN and Directed Diffusion allows negotiation between nodes concerning what data to forward to eliminate redundancy. Interest (REQ) and Data (DATA) caches maintained at each node.

Directed diffusion differs from SPIN in following aspects.

- Directed diffusion issues on demand data queries as the BS send queries to the sensor nodes by flooding some tasks. In SPIN, however, sensors advertise the availability of data allowing interested nodes to query that data.

- Second, all communication in directed diffusion is neighbor-to-neighbor with each node having the capability of performing data aggregation and caching. Unlike SPIN, there is no need to maintain global network topology in directed diffusion. However, directed diffusion may not be applied to applications (e.g., environmental monitoring) that require continuous data delivery to the BS. This is because the query- driven on demand data model may not help in this regard. Moreover, matching data to queries might require some extra overhead at the sensor nodes.

- One advantage of Directed Diffusion over SPIN is that it consumes much less energy by having less traffic compared to flooding. Also in terms of scalability it is based on local interaction only and It uses the best path available so it has good latency bound. Moreover It is robust because of retransmission of interest and low data rate gradients.

- The disadvantage is Gradient setup phase is expensive. It is not energy aware as the best paths might be used too often. Need to retransmit interest and alternate path maintenance is needed

7. CONCLUSION

Wireless sensors are going to be the next big thing in the future. Routing of data conserving the energy and bandwidth will be of utmost importance. These two algorithms define the standards for the future. These both algorithms have the common objective of trying to extend the lifetime of the sensor network, while not compromising data delivery. Although these routing techniques look promising, there are still many challenges that need to be solved in sensor networks.

8. REFERENCE

[1] Negotiation-Based Protocols for Disseminating Information in Wireless Sensor Network by Jonna Kulil, Wendi Heinzelman, Hari Balakrishnan

[2] Directed Diffusion for Wireless Sensor Networking by Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, John Heidemann, and Fabio Silva

[3] Routing Survey by Jamal N. Al-Karaki Ahmed E. Kamal

MOBILE IP (INTERNET PROTOCOL)

Maitri Patel
Sabar Institute of Technology for Girls, Tajpur
maitri_148@yahoo.co.in

Mansi Patel
Sabar Institute of Technology for Girls, Tajpur
mansipatel1488@yahoo.co.in

Happy Patel
Sabar Institute of Technology for Girls, Tajpur
happy_pn88@yahoo.co.in

Abstract

Wireless networks are stretching their legs day by day. With the increasing no. of mobile users wireless technology has become inevitable. Wireless networking is the first step towards the mobile communication system. As for wireless networking we use certain protocols for the communication thus definitely we need protocols for mobile communication. These protocols as in wireless networks are called Mobile IP or Mobile Internet Protocol. Mobile Internet Protocol is a new recommended Internet protocol designed to support the mobility of a user (host). Host mobility is becoming important because of the recent blossoming of laptop computers and the high desire to have continuous network connectivity anywhere the host happens to be. The development of Mobile IP makes this possible.

1. Introduction

The exponential growth of the Internet and the inexorable increase in native computing power of laptop computers and other digital wireless data communication devices has brought the need for mobile networking into sharp focus. As network services proliferate and become available ubiquitously, every network device will take advantage of mobile networking technology to offer maximum flexibility to the customers needing those devices. To understand the contrast between the current realities of IP connectivity and future possibilities, consider the transition toward mobility that has occurred in telephony over the past 20 years. An analogous transition in the domain of networking, from dependence on fixed points of attachment to the flexibility afforded by mobility, has just begun. As PDAs and the next generation of data-ready cellular phones become more widely deployed, a greater degree of connectivity is almost becoming a necessity for the business user on the go. Also called as Wireless Internet Protocol, Wireless IP. [2]

a) Motivation for Mobile IP Design:

The IP address of a host consists of two parts:

1) The higher order bits of the address determine the network on which the host resides.

2) The remaining low-order bits determine the host number.

IP decides the next-hop by determining the network information from the destination IP address of the packet. On the other hand, higher level layers like TCP maintain information about connections that are indexed by a quadruplet containing the IP addresses of both the endpoints and the port numbers. Thus, while trying to support mobility on the Internet under the existing protocol suite, we are faced with two mutually conflicting requirements: (1) a mobile node has to change its IP address whenever it changes its point of attachment, so that packets destined to the node are routed correctly, (2) to maintain existing TCP connections, the mobile node has to keep its IP address the same. Changing the IP address will cause the connection to be disrupted and lost.

Mobile IP, the standard proposed by IETF, is designed to solve the problem by allowing each mobile node to have two IP addresses and by transparently maintaining the binding between the two addresses. One of the IP addresses is the permanent home address that is assigned at the home network and is used to identify communication endpoints. The other is a temporary care-of address that represents the current location of the host. The main goals of Mobile IP are to make mobility transparent to the higher level protocols and to make minimum changes to the existing Internet infrastructure.

Mobile IP uses two IP addresses:
A fixed home address and a care-of address
that changes at each new point of attachment.

b) Overview of the Protocol:

Mobile IP supports mobility by transparently binding the home address of the mobile node with its care-of address. This mobility binding is maintained by some specialized routers known as mobility agents. Mobility agents are of two types - home agents and foreign agents. The home agent, a designated router in the home network of the mobile node, maintains the mobility binding in a mobility binding table where each entry is identified by the tuple <permanent home address, temporary care-of address, association lifetime>. Figure shows a mobility binding table. The purpose of this table is to map a mobile node's home address with its care-of address and forward packets accordingly.

Home Address	Care-of Address	Lifetime (in sec)
131.193.171.4	128.172.23.78	200
131.193.171.2	119.123.56.78	150

Figure: Mobility Binding Table

Foreign agents are specialized routers on the foreign network where the mobile node is currently visiting. The foreign agent maintains a visitor list which contains information about the mobile nodes currently visiting that network. Each entry in the visitor list is identified by the tuple: < permanent home address, home agent address, media address of the mobile node, association lifetime>. Figure shows an instance of a visitor list.

Home Address	Home Agent Address	Media Address	Lifetime (in s)
131.193.44.14	131.193.44.7	00-60-08-95-66-E1	150
131.193.33.19	131.193.33.1	00-60-08-68-A2-56	200

Figure: Visitor List

In a typical scenario, the care-of address of a mobile node is the foreign agent's IP address. There can be another kind of care-of address, known as collocated care-of address, which is usually obtained by some external address assignment mechanism. [1]

c) Terminology:

Mobile IP introduces the following new functional entities:

Mobile node – A mobile node is a host or a router that changes its point of attachment from one network or sub network to another. A mobile node may change its location without changing its IP address. It may continue to communicate with other Internet nodes at any location using its (constant) IP address, assuming link-layer connectivity to a point of attachment is available.

Home agent – A home agent is a router on a mobile node's home network that tunnels datagrams for delivery to the mobile node when it is away from home and maintains current location information for the mobile node.

Foreign agent – A foreign agent is a router on a mobile node's visited network that provides routing services to the mobile node while registered. The foreign agent detunnels and delivers datagram to the mobile node that were tunneled by the mobile node's home agent. The foreign agent may always be selected as a default router by registered mobile nodes. A mobile node is given a long term IP address on a home network. When away from its home network, a care-of address is associated with the mobile node and reflects the mobile node's current point of attachment. The mobile node uses its home address as the source address of all IP datagram that it sends, except during registration if it happens to acquire another IP address. [4]

d) Relationship of the Components of the Mobile IP:

The Mobile Node is a device such as a cell phone, personal digital assistant, or laptop whose software enables network roaming capabilities.

The Home Agent is a router on the home network serving as the anchor point for communication with

the Mobile Node; it tunnels packets from a device on the Internet, called a Correspondent Node, to the roaming Mobile Node. (A tunnel is established between the Home Agent and a reachable point for the Mobile Node in the foreign network.)

The Foreign Agent is a router that may function as the point of attachment for the Mobile Node when it roams to a foreign network, delivering packets from the Home Agent to the Mobile Node.

The care-of address is the termination point of the tunnel toward the Mobile Node when it is on a foreign network. The Home Agent maintains an association between the home IP address of the Mobile Node and its care-of address, which is the current location of the Mobile Node on the foreign or visited network.

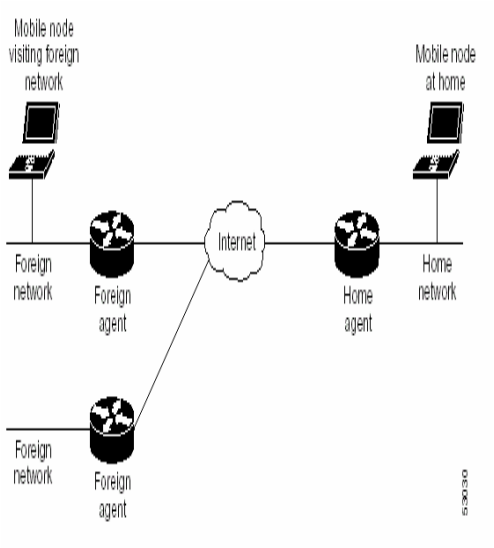


Figure: Mobile IP Component and Relationship

2. Working:

The basic Mobile IP protocol has four distinct stages. These are:

a) Agent Discovery: Agent Discovery consists of the following steps:

1. Mobility agents advertise their presence by periodically broadcasting Agent Advertisement messages. An Agent Advertisement message lists one or more care-of addresses and a flag indicating whether it is a home agent or a foreign agent.
2. The mobile node receiving the Agent Advertisement message observes whether the message is from its own home agent and determines whether it is on the home network or a foreign network.
3. If a mobile node does not wish to wait for the periodic advertisement, it can send out Agent Solicitation messages that will be responded by a mobility agent.

b) Registration: Registration consists of the following steps:

1. If a mobile node discovers that it is on the home network, it operates without any mobility services.
2. If the mobile node is on a new network, it registers with the foreign agent by sending a Registration Request message which includes the permanent IP address of the mobile host and the IP address of its home agent.
3. The foreign agent in turn performs the registration process on behalf of the mobile host by sending a Registration Request containing the permanent IP address of the mobile node and the IP address of the foreign agent to the home agent.
4. When the home agent receives the Registration Request, it updates the mobility binding by associating the care-of address of the mobile node with its home address.
5. The home agent then sends an acknowledgement to the foreign agent.
6. The foreign agent in turn updates its visitor list by inserting the entry for the mobile node and relays the reply to the mobile node.

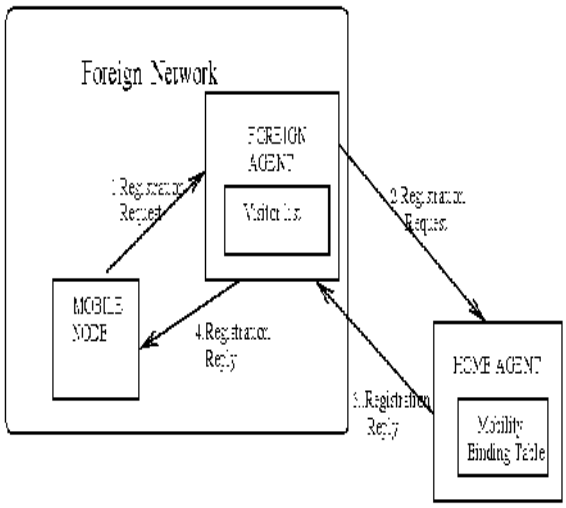


Figure: Registration process in Mobile IP

C. In Service: This stage can be subdivided into the following steps:

1. When a correspondent node wants to communicate with the mobile node, it sends an IP packet addressed to the permanent IP address of the mobile node.
2. The home agent intercepts this packet and consults the mobility binding table to find out if the mobile node is currently visiting any other network.
3. The home agent finds out the mobile node's care-of address and constructs a new IP header that contains the mobile node's care-of address as the destination IP address. The original IP packet is put into the payload of this IP packet. It then sends the packet. This process of encapsulating one IP packet into the payload of another is known as IP-within-IP encapsulation or tunneling.
4. When the encapsulated packet reaches the mobile node's current network, the foreign agent decapsulates the packet and finds out the mobile node's home address. It then consults the visitor list to see if it has an entry for that mobile node.
5. If there is an entry for the mobile node on the visitor list, the foreign agent retrieves the corresponding media address and relays it to the mobile node.
6. When the mobile node wants to send a message to a correspondent node, it forwards the packet to the foreign agent, which in turn relays the packet to the correspondent node using normal IP routing.
7. The foreign agent continues serving the mobile node until the granted lifetime expires. If the mobile node wants to continue the service, it has to reissue the Registration Request.

4. Deregistration: If a mobile node wants to drop its care-of address, it has to deregister with its home agent. It achieves this by sending a Registration Request with the lifetime set to zero. There is no need for deregistering with the foreign agent as registration automatically expires when lifetime becomes zero. However if the mobile node visits a new network, the old foreign network does not know the new care-of address of the mobile node. Thus datagrams already forwarded by the home agent to the old foreign agent of the mobile node are lost. [1]

a) Packet Transmission Example:

Figure illustrates the basic operation when a mobile host (MH1) within range of a foreign agent FA1, having a home agent HA1, wants to communicate with another mobile host (MH2) within range of a foreign agent FA2, having a home agent HA2. The following operations are shown in Figure:

1. MH1 and MH2 both register with their foreign agents and notify their home agents of their new bindings.
2. Suppose that MH1 wants to send a packet to MH2 and MH1 does not know the care-of address of MH2. MH1 transmits the packet relying on existing routing protocols, through FA1. The packet is received by MH2's home agent, HA2, which tunnels the packet to MH2's foreign agent FA2. When FA2 receives the tunneled packet, it decapsulates it and delivers it locally.
3. When HA2 receives and the tunnels the packet, it also sends to the source (here, MH1) an IMHP Binding Notify packet containing MH2's binding, as MH1 seems not to have a binding cached for MH2.
4. MH1 can transmit future packets for MH2 by tunneling them directly to MH2's current foreign agent, FA2. A close to optimum route is thus established.

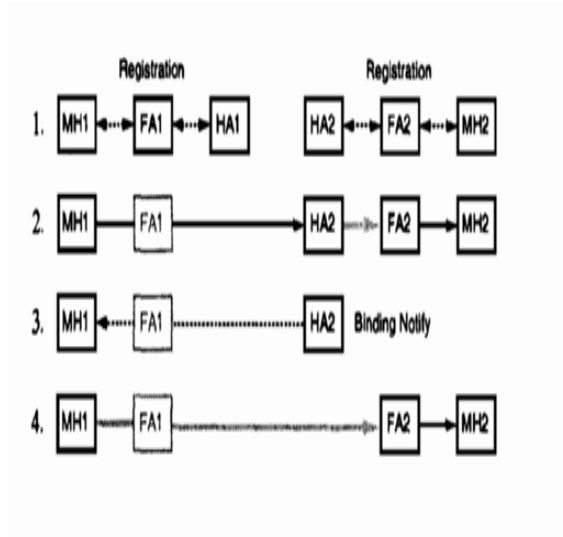


Figure: Mobile host to mobile host communication.

Figure describes the same situation which was the case at the end of the previous example. It describes what happens when MH2 migrates to a new location.

The following operations are shown in Figure:

1. MH2 detects that it is connected to a new network. It registers with a new foreign agent, FA3, and notifies MH2's home agent (HA2).
2. MH2's previous foreign agent, FA2, is also notified of MH2's new binding.
3. Suppose that MH1 wants to send a packet to where it believes MH2 to be located (FA2). FA2 forwards the packet to FA3. FA3 decapsulates the packet and delivers it locally to MH2.
4. FA2 recognizes that MH1 must have an old binding for MH2, since otherwise MH1 would not have tunneled the packet to FA2. FA2 thus sends MH1 a Binding Notify packet notifying it of MH2's new binding at FA3.
5. Future packets to MH2 are tunneled directly to FA3.

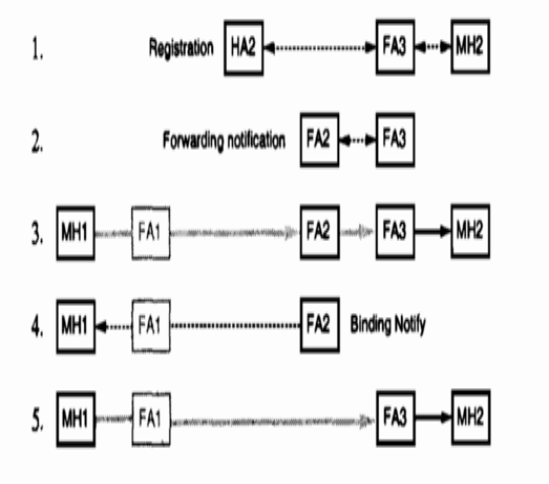


Figure: Mobile host movement

It should be noted that all the binding notifies may need to be authenticated for security reasons. [3]

3. Security:

Security is very important in Mobile IP as mobile nodes are often connected to the Internet via wireless links which are very vulnerable to security attacks. For example, during registration procedure the home agent should be convinced that it is getting authentic Registration Request from a mobile node and not receiving information from a bogus node. Mobile IP solves this problem by specifying a security association between the home agent and the mobile node. This security association is at present manually configured. Every registration message should contain a mobile node-home agent authentication extension which contains a Security Parameters Index (SPI) followed by an authenticator. The SPI is an index into the mobility security association and it defines the security context (i.e., the algorithm and the secret) used to compute and check the authenticator. Two methods are used to generate the unique data.

1)Timestamps - The node generating the message inserts the time-of-day, and the node receiving the message checks whether it is sufficiently close to its time-of-day.

2) Nonces - Node A generate a new random number in every message to node B, and checks whether node B returns the same number in its next message to node A. Both messages use an authentication code to protect against alteration by an outsider. Node B can also generate random numbers and use them in its messages. [5]

4. Ongoing Work:

Work is also continuing to refine and extend the protocol within the academic and commercial communities and within the IETF. This section surveys the state of implementation of Mobile IP and speculates on a possible timetable for deployment.

1) Routing Inefficiencies
2) Security Issues
3) Ingress Filtering
4) Issues In IP Addressing
5) Competition From Other Protocol
6) Congestion

a) Mobile IPv6:

The next version of IP, IPv6 is designed to be an evolutionary step from IPv4.IPv6 addresses are 128 bits long. Mobility support in IPv6 solves many of the problems of basic Mobile IP. Some advantages of Mobile IPv6 over Mobile IPv4 are:

1) Route Optimization is built as a fundamental part of Mobile IPv6 unlike Mobile IPv4 where it is an optional set of extensions that may not be supported by all nodes.
2) Foreign Agents are not needed in Mobile IPv6. The enhanced features of IPv6 like Neighbor Discovery and Address Auto configuration enable mobile nodes to function in any location without the services of any special router in that location.

3) In Mobile IPv4, when a mobile node communicates with a correspondent node, it puts its home address as the source address of the packet. Thus “ingress filtering routers "used to filter out the packets as the source address of the packet is different from the network from which the packet originated. This problem is tackled in Mobile IPv6 by putting the care-of address as the source address and having a Home Address Destination option, allowing the use of the care-of address to be transparent over the IP layer.

The biggest difference between IPv6 and IPv4 is that all IPv6 nodes are expected to implement strong authentication and encryption features to improve Internet security. This affords a major simplification for IPv6 mobility support, since all authentication procedures can be assumed to exist when needed and do not have to be specified in the Mobile IPv6 protocol. Even with the security features in IPv6, however, the current working group draft for IPv6 mobility support specifies the use of authentication procedures as infrequently as possible. [6]

5. Conclusion:

The development of Mobile IP has been going on for over half a decade. By today, the proposals have come to a point where there are no widely known drawbacks. Mobile IP is at its best in a large network. Roaming in an area where the transceivers cover only a very small geographic area, the mobility can be achieved by a simpler manner.

6. References:

[1] http://www.acm.org/crossroads/xrds7-2/mobileip.html

[2] http://en.wikipedia.org/wiki/Mobile_ip

[3] http://www.tml.tkk.fi/Opinnot/Tik-111.550/1999/Esitelmat/MobileIP/Mobip.html

[4]www.ee.oulu.fi/~skidi/teaching/mobile_and_ubiquitous_multimedia_2002/Perkins.pdf

[5] Mobile IP Technology and Applications by Madhavi W. Chandra

[6] Mobil IP: Design Principles and Practices by Charles E. Perkins

Graphics Processing Unit With Graphics API

Darshana Mistry

Computer Engineering Department, Gandhinagar Institute Of Technology
Gandhinagar, India
darhshana.mistry@git.org.in,

***Abstract*— Graphics Processing Unit(GPU) is dedicated to graphics rendering device for a personal computer, workstation or game console. GPU implements a number of graphics primitive operations in way that makes running them much faster than drawing directly to the screen with the host CPU with different graphics language/API as Cg, OpenGL, or DirectX. DirectX is method because of it support multimedia and 3D graphics.**

***Keywords*— Graphics Processing Unit, GPU, Cg, OpenGL, DirectX**

I. INTRODUCTION

GRAPHICS processing unit or GPU (also occasionally called visual processing unit or VPU) is a dedicated graphics rendering device for a personal computer, workstation, or game console. Modern GPUs are very efficient at manipulating and displaying computer graphics, and their highly parallel structure makes them more effective than general-purpose CPUs for a range of complex algorithms. A GPU can sit on top of a video card, or it can be integrated directly into the motherboard.

The technical definition :- "a single chip processor with integrated transform, lighting, triangle setup/ clipping, and rendering engines that is capable of processing a minimum of 10 million polygons per second".

In 1980, the Commodore Amiga was the first mass-market computer to include a blitter in its video hardware, and IBM's 8514 graphics system was one of the first PC video cards to implement 2D primitives in hardware. By the early 1990s, the rise of Microsoft Windows sparked a surge of interest in high-speed, high-resolution 2D bitmapped graphics (which had previously been the domain of Unix workstations and the Apple Macintosh). For the PC market, the dominance of Windows meant PC graphics vendors could now focus development effort on a single programming interface, Graphics Device Interface (GDI).

Direct3D 5.0 was the first version of the burgeoning API to really dominate the gaming market and stomp out many of the hardware-specific interfaces. With the advent of the OpenGL API and similar functionality in DirectX, GPUs added programmable shading to their capabilities. Each pixel could now be processed by a short program that could include additional image textures as inputs, and each geometric vertex could likewise be processed by a short program before it was projected onto the screen. NVIDIA was first to produce a chip capable of programmable shading, the GeForce 3 (core named NV20).

Today, parallel GPUs have begun making computational inroads against the CPU, and a subfield of research, dubbed GPGPU for General Purpose Computing on GPU, has found its way into fields as diverse as oil exploration, scientific image processing, and even stock options pricing determination. There is increased pressure on GPU manufacturers from "GPGPU users" to improve hardware design, usually focusing on adding more flexibility to the programming model. The current dominators of the market are AMD (manufacturers of the ATI Radeon graphics chip line) and NVIDIA (manufacturers of the NVIDIA Geforce and NVIDIA Quadra graphics chip line.)

II. GRAPHICS PIPELINE

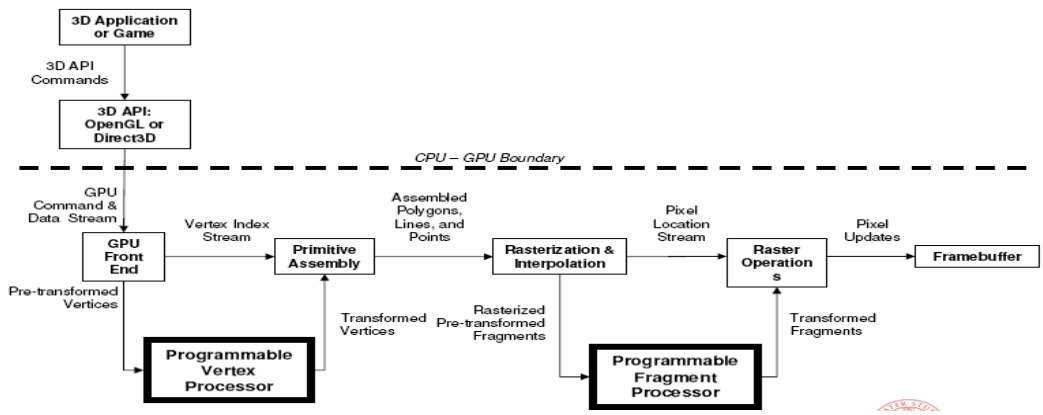


Fig. 1. GPU Pipeline

Process of GPU Pipeline as below:

1. The CPU sends instructions (compiled shading language programs) and geometry data to the graphics processing unit, located on the graphics card
2. Within the vertex shader, the geometry is transformed and lighting calculations are performed.
3. If a geometry shader is in the graphic processing unit, some changes of the geometries in the scene are performed.
4. The calculated geometry is triangulated (subdivided into triangles). Triangles are transformed into pixel quads (one pixel quad is a 2×2 pixel primitive).This is only true for Z-Buffering visibility style tests, other hidden surface removal may be done at the application level.
5. In addition to the 3D hardware, today's GPUs include basic 2D acceleration and frame buffer capabilities (usually with a VGA compatibility mode).

III. GRAPHICS API

Direct3D, OpenGL etc. are used Graphics API. High level Shading language (HLSL), Cg (c for Graphics), GLSL (Graphics Language Shading Language) etc are used to improve Graphics.

A. Cg (C for Graphics)

It is a high-level shading language developed by Nvidia in close collaboration with Microsoft for programming vertex and pixel shaders. It is very similar to Microsoft's HLSL. Cg is based on the C programming language and although they share the same syntax, some features of C were modified and new data types were added to make Cg more suitable for programming graphics processing units. This language is only suitable for GPU programming and it does not replace a general programming language. The Cg compiler outputs DirectX or OpenGL shader programs.

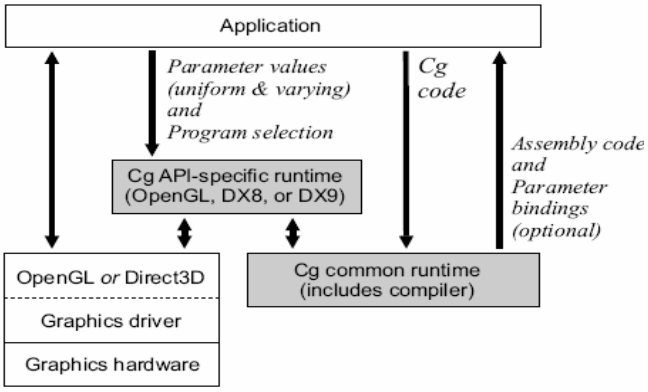


Fig. 2. Cg System Architecture

1) Data types and Operators

Cg has six basic data types, some of them are the same as in C, others are especially added for GPU programming, these types are:

float - a 32bit floating point number
half - a 16bit floating point number
int - a 32bit integer
fixed - a 12bit fixed point number
bool - a boolean variable
sampler - represents a texture object

Cg supports a wide range of operators, including the common arithmetic operators from C, the equivalent arithmetic operators for vector and matrix data types, and the common logical operators. Control statements as like of C language. Some function as like C language and some are special for graphics as like texture mapping functions tex1D and tex2D.

2) The Cg Runtime Library

Cg programs are merely vertex and pixel shaders, and they need supporting programs that handle the rest of the rendering process. Cg can be used with two APIs: OpenGL or DirectX. Each has its own set of Cg functions to communicate with the Cg program, like setting the current Cg shader, passing parameters, and such tasks.

In addition to being able to compile Cg source to assembly code, the Cg runtime also has the ability to compile shaders during execution of the supporting program. This allows the runtime to compile the shader using the latest optimizations available for hardware that the program is currently executing on. However, this technique requires that the source code for the shader be available in plain text to the compiler, allowing the user of the program to access the source-code for the shader, Some developers view this as a major drawback of this technique.

To avoid exposing the source code of the shader, and still maintain some of the hardware specific optimizations, the concept of profiles was developed. Shaders can be compiled to suit different graphics hardware platforms (according to profiles). When executing the supporting program, the best/most optimized shader is loaded according to its profile. For instance there might be a profile for a graphics card that supports complex pixel shaders, and another profile for one that supports only minimal pixel shaders. By creating a pixel shader for each of these profiles a supporting program enlarges the number of supported hardware platforms without sacrificing picture quality on powerful systems.

The application must supply the modelViewProjection matrix and the brightness scalar, typically by using the Cg runtime library's API. The POSITION, COLOR, and

TEXCOORD0 indenters following the objectPosition, color, and decalCoord parameters specify how these parameters are bound to API resources. In OpenGL, glVertex commands feed POSITION; glColor commands feed COLOR; and glMultiTexCoord commands feed TEXCOORDn. The out modi_er indicates that clipPosition, oColor, and oDecalCoord parameters are output by the program. The identi_er following the colon after each of these parameters speci_es how the output is fed to the primitive assembly and rasterization stages of the graphics pipeline.

B. OpenGL (Open Graphics Library)

OpenGL (Open Graphics Library) is a standard specification defining a cross-language cross-platform API for writing applications that produce 2D and 3D computer graphics. The interface consists of over 250 different function calls which can be used to draw complex three-dimensional scenes from simple primitives. OpenGL was developed by Silicon Graphics Inc. (SGI) in 1992 and is widely used in CAD, virtual reality, scientific visualization, information visualization, and flight simulation. It is also used in video games, where it competes with Direct3D on Microsoft Windows platforms. Latest release is OpenGL 2.1[1] and [3].

1) Design

OpenGL serves two main purposes:

- To hide the complexities of interfacing with different 3D accelerators, by presenting the programmer with a single, uniform API.
- To hide the differing capabilities of hardware platforms, by requiring that all implementations support the full OpenGL feature set (using software emulation if necessary).

OpenGL's basic operation is to accept primitives such as points, lines and polygons, and convert them into pixels. This is done by a graphics pipeline known as the OpenGL state machine. Most OpenGL commands either issue primitives to the graphics pipeline, or configure how the pipeline processes these primitives.

OpenGL is a low-level, procedural API, requiring the programmer to dictate the exact steps required to render a scene. This contrasts with descriptive (aka scene graph or retained mode) APIs, where a programmer only needs to describe a scene and can let the library manage the details of rendering it. OpenGL's low-level design requires programmers to have a good knowledge of the graphics pipeline, but also gives a certain amount of freedom to implement novel rendering algorithms.

2) Higher Level Functionality

OpenGL was designed to be graphic output-only: it provides only rendering functions. The core API has no concept of windowing systems, audio, printing to the screen, keyboard/mouse or other input devices. While this seems restrictive at first, it allows the code that does the rendering to be completely independent of the operating system it is running on, allowing cross-platform development. However some integration with the native windowing system is required to allow clean interaction with the host system. This is performed through the following add-on APIs:

- GLX – X11 (including network transparency)
- WGL – Microsoft Windows
- CGL – Mac OS X. Better integration with Mac OS X's application frameworks is provided by APIs layered on top of CGL: AGL for Carbon and NSOpenGL for Cocoa.
- Additionally, GLUT, SDL and the GLFW libraries provide functionality for basic windowing using OpenGL, in a portable manner.

C. DirectX

Microsoft DirectX is a collection of API for handling tasks related to multimedia, especially game programming and video, on Microsoft platforms. Originally, the names of these APIs all began with Direct, such as Direct3D, DirectDraw, DirectMusic, DirectPlay, DirectSound, and so forth. DirectX, then, was the generic term for all of these Direct-something APIs, and that term became the name of the collection. Direct3D is also used by other software applications for visualization and graphics tasks, most notably among the engineering sector for CAD/CAM, because of its ability to quickly render high-quality 3D graphics using DirectX-compatible graphics hardware.

The DirectX software development kit consists of runtime libraries in redistributable binary form, along with accompanying documentation and headers for use in coding. Originally, the runtimes were only installed by games or explicitly by the user. The SDK is available as a free download. While the runtimes are proprietary, closed-source software, source code is provided for most of the SDK samples.

The latest versions of Direct3D, namely, Direct3D 10 and Direct3D 9Ex, are exclusive to Windows Vista. The reasons, as Microsoft claims, is that there were extensive changes in the Windows graphics architecture, and in particular the introduction of the Windows Display Driver Model. Critics argue that there may be possible commercial motives as well. This redesign of the graphics infrastructure for Windows Vista supports virtualizing graphics hardware to multiple applications and services such as the Desktop Window Manager, in contrast to the exclusive access afforded to DirectX applications on Windows XP. Both Direct3D 9Ex and Direct3D 10 rely on the WDDM infrastructure and WDDM drivers[1].

1) Component:-

The components comprising DirectX are

- DirectX Graphics, which consists of several APIs: DirectDraw: for drawing 2D Graphics

(raster graphics). Now deprecated, though still in use by a number of games.
    - Direct3D(D3D): for drawing 3D graphics.
    - DXGI: for enumerating adapters and monitors and managing swap chains for Direct3D 10 and up.
- DirectInput: for interfacing with input devices including keyboards, mice, joysticks, or other game controllers. Deprecated after version 8 in favour of XInput for Xbox360 controllers or standard WM INPUT window message processing for keyboard and mouse input.

- DirectPlay: for communication over a local-area or wide-area network. Deprecated after version 8.

- DirectSound: for the playback and recording of waveform sounds. DirectSound3D (DS3D): for the playback of 3D sounds

- DirectMusic: for playback of soundtracks authored in DirectMusic Producer.

- DirectX Media: comprising DirectAnimation for 2D web animation, DirectShow for multimedia playback and streaming media, DirectX Transform for web interactivity, and Direct3D Retained Mode for higher level 3D graphics. DirectShow contains DirectX plug-in for audio signal processing and DirectX Video Acceleration for accelerated video playback.

- DirectX Media Objects: support for streaming objects such as encoders, decoders, and effects.

- DirectSetup: for the installation of DirectX components. Not a game API per se.

2)DirectX 10

- A major update to DirectX API, DirectX 10 ships with and is only available with Windows Vista; previous versions of Windows are not able to run DirectX 10-exclusive applications. Changes for DirectX 10 were extensive, but only Direct3D featured a major overhaul of the API. XInput, DirectSound was deprecated in favour of XACT and lost support for hardware accelerated audio, since Vista audio stack renders sound in software on the CPU. The DirectPlay DPLAY.DLL was also removed and was replaced with dplayx.dll; games that rely on this DLL must duplicate it and rename it to dplay.dll.

  In order to achieve backwards compatibility, DirectX in Windows Vista contains several versions of Direct3D:

- **Direct3D 9:** emulates Direct3D 9 behaviour as it was on Windows XP. Details and advantages of Vista's Windows Display Driver Model are hidden from the application if WDDM drivers are installed. This is the only. API available if there are only XP graphic drivers (XPDM) installed, after an upgrade to Vista for example
- **Direct3D 9Ex:-** (known internally during Windows Vista development as 9.0L or 9.L, the L referring to Longhorn, the codename for Windows Vista): allows full access to the new capabilities of WDDM (if WDDM drivers are installed) while maintaining compatibility for existing Direct3D applications. The Windows Aero user interface relies on D3D 9Ex.
- **Direct3D 10**: Designed around the new driver model in Windows Vista and featuring a number of improvements to rendering capabilities and flexibility, including shader Model 4.
- Direct3D 10.1 is an incremental update of Direct3D 10.0 which will be shipped with, and require, Windows Vista Service Pack 1 in February 2008. This release mainly sets a few more image quality standards for graphics vendors, while giving developers more control over image quality. It also requires a whole new set of requirements, including Shader Model 4.1 and 32-bit floating-point operations. DirectX 10.1 still fully supports DirectX 10 hardware, but to be able to support all the features you need a card with the required hardware. As of March 5, 2008, only the ATI Radeon 3000 series GPUs are compliant.

IV. Conclusion

GPUs are designed for graphics. GPUs are fast, powerful, hard to program, highly parallel tasks than CPU. GPUs process independent vertices & fragments. Hard to keep up with these features, high-level languages needed for programming. Using of Cg, OpenGL, or DirectX. DirectX are best method because of it support to multimedia and 3D graphics.

Acknowledgment

I am very thankful to Dr S.N. Pradhan who is a professor of Computer Department, Nirma University for guiding throughout this major project part and providing feedback to improve my work. I heartily thankful to her for his time to time suggestion and clarity of concepts of topic that helped me a lot during this study.

References

[1] Electronic Publication: Graphics Processing Unit

[2] J. Talton, D. Fitzpatrick, "Teaching graphics with the OpenGL Shading language", Proceedings of the 38th SIGCSE technical symposium on Computer science education, March 2007

[3] W. Mark, R. S.Glanvilley ,K. Akeleyy, M. Kilgardy, "The Cg tutorial, the definitive guide to programmble real-time graphics, March 2003

Darshana Mistry(ISTE-LM'10). She became a life member ship of ISTE in 2010. Her date of birth is 1st May 1981 in Bharuch. She completed H.S.C from G.S.E.B(Gujarat) in 1998 with 83.67%, B.E.Computer Engineer from S.P.University(V.V.Nagar,Gujarat) with 5.78 CPI in 2002, M.Tech(CSE) from Nirma University(Ahmedabad,Gujarat) with 7.31 CPI in 2009.

She is working as **HOD of CE dept., Asst. Professor in Gandhinagar Institute of Technology college, Gandhinagar.** She has 7 years experience. She published 1 state level(Gandhinagar, Gujarat, GIT journal,2010), 1national level(Ahmedabad,Gujarat, NUCONE-09-Nirma University,2009), 1 International level paper(V.V.Nagar,Gujarat, ICSSA-09,GCET,2009),1 paper in International Journal in Computer Science and Engineering Volume II, Issue-IV, ISSN 2905-2909, December 2010, 1 paper in International Journal in Computer science and Engineering Technology will be published in March 2011.

Performance Comparison of AODV/DSR On-demand Routing Protocols for Mobile Ad Hoc Networks (April – 2011)

Mr. Margil P. Shah
Computer Department
Gandhinagar Institute of Technology
margil.shah@git.org.in

Ms. Bhumi S. Jani
Information Technology Department
Kalol Institute of Technology and Research Centre
bhumi_mina@yahoo.co.in

Abstract - Ad hoc networks are characterized by multihop wireless connectivity, frequently changing network topology and the need for efficient dynamic routing protocols.

We compare the performance of two prominent on demand routing protocols for mobile ad hoc networks—Dynamic Source Routing (DSR) and Ad Hoc On-Demand Distance Vector Routing (AODV). We demonstrate that even though DSR and AODV share a similar on-demand behavior, the differences in the protocol mechanics can lead to significant performance differentials. The AODV out-perform DSR in the normal situation but in the constrained situation DSR out-performs AODV, the degradation is as severe as (30%) in AODV whereas DSR degrades marginally (10%) as observed through simulation.

Keywords- Ad hoc networks, routing protocols, mobile networks, wireless networks, simulation, performance evaluation

1. INTRODUCTION

Wireless cellular systems have been in use since 1980s. We have seen their evolutions to first, second and third generation's wireless systems. These systems work with the support of a centralized supporting structure such as an access point. The wireless users can be connected with the wireless system by the help of these access points, when they move or roam from one place to the other.

The adaptability of wireless systems is limited by the presence of a fixed supporting coordinate. It means that the technology cannot work efficiently in that places where there is no permanent infrastructure. Easy and fast deployment of wireless networks will be expected by the future generation wireless systems. This fast network deployment is not possible with the existing structure of present wireless systems.

Recent advancements such as Bluetooth introduced a fresh type of wireless systems which is frequently known as mobile ad-hoc networks. Mobile ad-hoc networks or "short live" networks control in the nonexistence of permanent infrastructure. Mobile ad hoc net work offers quick and horizontal network deployment in conditions where it is not possible otherwise. Ad-hoc is a Latin word, which means "for this or for this only."

Mobile ad hoc network is an autonomous system of mobile nodes connected by wireless links; each node operates as an end system and a router for all other nodes in the network. A wireless network is a growing new technology that will allow users to access services and information electronically, irrespective of their geographic position.

Wireless networks can be classified in two types: - infrastructure network and infrastructure less (ad hoc) networks. Infrastructure network consists of a network with fixed and wired gateways. A mobile host interacts with a bridge in the network (called base station) within its communication radius. The mobile unit can move geographically while it is communicating. When it goes out of range of one base station, it connects with new base station and starts communicating through it. This is called handoff. In this approach the base stations are fixed.

A Mobile ad hoc network is a group of wireless mobile computers (or nodes); in which nodes collaborate by forwarding packets for each other to allow them to communicate outside range of direct

wireless transmission. Ad hoc networks require no centralized administration or fixed network infrastructure such as base stations or access points, and can be quickly and inexpensively set up as needed.

A MANET is an autonomous group of mobile users that communicate over reasonably slow wireless links. The network topology may vary rapidly and unpredictably over time, because the nodes are mobile. The network is decentralized, where all network activity, including discovering the topology and delivering messages must be executed by the nodes themselves. Hence routing functionality will have to be incorporated into the mobile nodes.

MANET is a kind of wireless ad-hoc network and it is a self-configuring network of mobile routers (and associated hosts) connected by wireless link the union of which forms an arbitrary topology. The routers, the participating nodes act as router, are free to move randomly and manage themselves arbitrarily; thus, the network's wireless topology may change rapidly and unpredictably. Such a network may operate in a standalone fashion, or may be connected to the larger Internet.

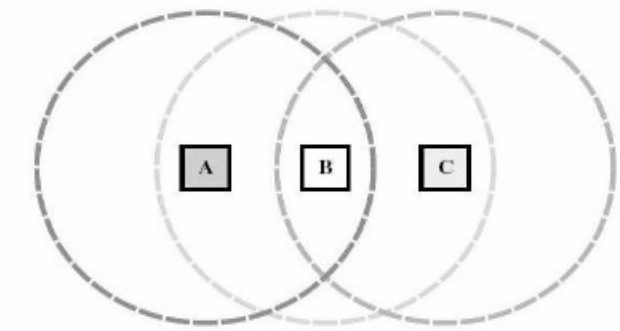


Figure 1.1 Example of a simple ad-hoc network with three participating nodes

Mobile ad hoc network is a collection of independent mobile nodes that can communicate to each other via radio waves. The mobile nodes can directly communicate to those nodes that are in radio range of each other, whereas others nodes need the help of intermediate nodes to route their packets. These networks are fully distributed, and can work at any place without the aid of any infrastructure. This property makes these networks highly robust.

In Figure 1.1 nodes A and C must discover the route through B in order to communicate. are not in direct transmission range of each other, since A‘s circle does not cover C.

1.1. Characteristics of MANET

Mobile ad hoc network nodes are furnished with wireless transmitters and receivers using antennas, which may be highly directional (point-to-point), omnidirectional (broadcast), probably steerable, or some combination thereof. At a given point in time, depending on positions of nodes, their transmitter and receiver coverage patterns, communication power levels and co-channel interference levels, a wireless connectivity in the form of a random, multihop graph or "ad hoc" network exists among the nodes. This ad hoc topology may modify with time as the nodes move or adjust their transmission and reception parameters.

The characteristics of these networks are summarized as follows:

- Communication via wireless means.
- Nodes can perform the roles of both hosts and routers.
- Bandwidth-constrained, variable capacity links.
- Energy-constrained Operation.
- Limited Physical Security.
- Dynamic network topology.
- Frequent routing updates.

1.2. Advantages of MANET

The following are the advantages of MANET:

- They provide access to information and services regardless of geographic position.
- These networks can be set up at any place and time.

1.3. Disadvantages of MANET

Some of the disadvantages of MANETs are as follows:

- Limited resources and physical security.
- Intrinsic mutual trust vulnerable to attacks.
- Lack of authorization facilities.
- Volatile network topology makes it hard to detect malicious nodes.
- Security protocols for wired networks cannot work for ad hoc networks.

1.4. Applications of MANET

Some of the applications of MANETs are as follows:

- Military or police exercises.
- Disaster relief operations.
- Mine cite operations.
- Urgent Business meetings.

2. Routing in MANETs

A routing protocol is the mechanism by which user traffic is directed and transported through the network from the source node to the destination node. Objectives

include maximizing network performance from the application point of view - application requirements- while minimizing the cost of network itself in accordance with its capacity. The application requirements are hop count, delay, throughput, loss rate, stability, jitter, cost; and the network capacity is a function of available resources that reside at each node and number of nodes in the network as well as its density, frequency of end-to-end connection (i.e. number of communication), frequency of topology changes (mobility rate). The four core basic routing functionalities for mobile ad hoc networks are:

Path generation: This generates paths according to the assembled and distributed state information of the network and of the application; assembling and distributing network and user traffic state information

Path selection: This selects appropriate paths based on network and application state information.

Data Forwarding: This forwards user traffic along the select route forwarding user traffic along the selected route.

Path Maintenance: Maintaining of the selected route. Consequently routing is bounded by traffic requirements, network capacity and the security requirements, as illustrated in Figure. 2.1

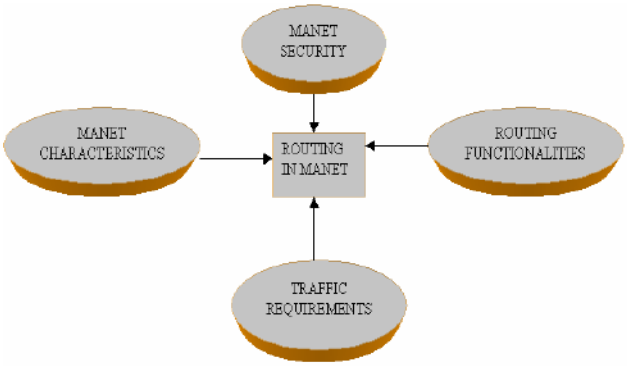


Figure 2.1 Routing in MANET

Due to its characteristics, other desirable features of ad hoc routing protocol include- fast route establishment, multiple routes selection, energy/bandwidth efficiency and fast adaptability to link changes. Almost all routing systems respond in some way to the changes in network and user traffic state. However, routing systems vary widely in the types of state changes to which they respond and the speed of their response. Routing states can be divided into three categories - Static, Quasi Static and Dynamic. Further, each of the three basic routing functions may be implemented in three ways- Centralized, Decentralized and Distributed.

On-Demand Routing (Reactive Protocols)

In reactive protocols, routes are determined when they are required by the source using a route discovery process. These protocols were designed to reduce the overhead encountered in proactive protocols by maintaining information for active routes only. This means that the routes are determined and maintained for the nodes that are required to send data to a particular destination. Route discovery usually occurs by flooding route request packets through the network. When a node with a route to the destination (or the destination itself) is reached a route reply is sent back to the source node using link reversal if the route request has traveled through the bi-directional links or by piggy-backing the route in a route reply packet via flooding. Therefore, the route discovery overhead (in the worst case scenario) will grow by O(N+M) when link reversal is possible and O(2N) for unidirectional links (where, N represents the total number of nodes and M represents the total number of nodes in the localized region).

Reactive protocols can be classified into two categories:

- Source routing, and
- Hop-by-Hop routing

In Source routed on-demand protocols, each data packets carry the complete source to destination address. Therefore, each intermediate node forwards these packets according to the information kept in the header of each packet. This means that the intermediate nodes do not need to maintain up-to-date routing information for each active route in order to forward the packet towards the destination. Furthermore, nodes do not need to maintain neighbor's connectivity through periodic beaconing messages. The major drawback with source routing protocols is that in large networks they do not perform well.

In hop-by-hop routing (also known as point-to-point routing), each data packet only carries the destination address and the next hop address. Therefore, each intermediate node in the path to the destination uses its routing table to forward each data packet towards the destination. The advantage of this strategy is that routes are adaptable to the dynamically changing environment of MANETs, since each node can update its routing table when they receiver fresher topology information and hence forward the data packets over fresher and better routes. The disadvantage of this

strategy is that each intermediate node must store and maintain routing information for each active route and each node may require being aware of their surrounding neighbors through the use of beaconing messages. This following section describes the three protocols along with their performance comparison. The performance metrics represent the worst-case scenario.

1) Ad hoc On-demand Distance Vector (AODV)

The AODV routing protocol is based on DSDV and DSR algorithm. It uses the periodic beaconing and sequence numbering procedure of DSDV and a similar route discovery procedure as in DSR. However, there are two major differences between DSR and AODV. The most distinguishing difference is that in DSR each packet carries full routing information, whereas in AODV the packets carry the destination address. This means that AODV has potentially less routing overheads than DSR. The other difference is that the route replies in DSR carry the address of every node along the route, whereas in AODV the route replies only carry the destination IP address and the sequence number. The advantage of AODV is that it is adaptable to highly dynamic networks. However, node may experience large delays during route construction, and link failure may initiate another route discovery, which introduces extra delays and consumes more bandwidth as the size of the network increases.

Ad hoc On-Demand Distance Vector (AODV) routing is a routing protocol for mobile ad hoc networks and other wireless ad-hoc networks. It is jointly developed in Nokia Research Centre of University of California, Santa Barbara and University of Cincinnati by C. Perkins and S. Das. It is an on-demand and distance-vector routing protocol, meaning that a route is established by AODV from a destination only on demand[1,4].

AODV is capable of both unicast and multicast routing. It keeps these routes as long as they are desirable by the sources. Additionally, AODV creates trees which connect multicast group members. The trees are composed of the group members and the nodes needed to connect the members. The sequence numbers are used by AODV to ensure the freshness of routes. It is loop-free, self-starting, and scales to large numbers of mobile nodes. AODV defines three types of control messages for route maintenance:

RREQ- A route request message is transmitted by a node requiring a route to a node. As an optimization AODV uses an expanding ring technique when flooding these messages. Every RREQ carries a time to live (TTL) value that states for how many hops this message should be forwarded. This value is set to a predefined value at the first transmission and increased at retransmissions. Retransmissions occur if no replies are received. Data packets waiting to be transmitted (i.e. the packets that initiated the RREQ). Every node maintains two separate counters: a node sequence number and a broadcast_ id. The RREQ contains the following fields:-
Source address, broadcast ID, source sequence number, destination address, destination sequence number and hop count[2,3].

The pair <source address, broadcast ID> uniquely identifies a RREQ. Broadcast_id is incremented whenever the source issues a new RREQ.

RREP- A route reply message is unicasted back to the originator of a RREQ if the receiver is either the node using the requested address, or it has a valid route to the requested address. The reason one can unicast the message back, is that every route forwarding a RREQ caches a route back to the originator[2].

RERR- Nodes monitor the link status of next hops in active routes. When a link breakage in an active route is detected, a RERR message is used to notify other nodes of the loss of the link. In order to enable this reporting mechanism, each node keeps a ―precursor list", containing the IP address for each its neighbors that are likely to use it as a next hop towards each destination[3].

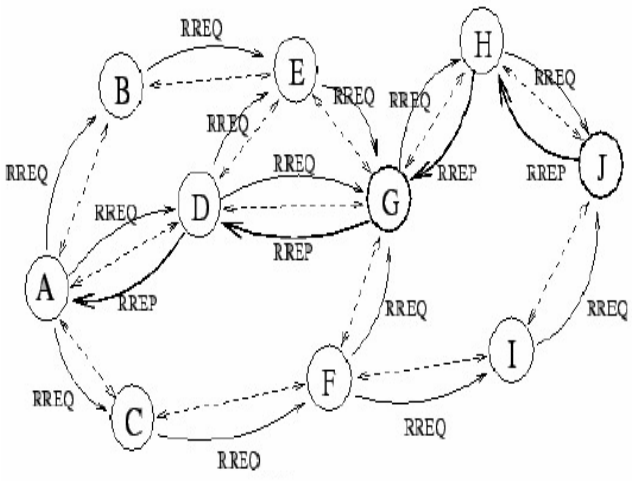


Figure 2.2: A possible path for a route replies if A wishes to find a route to J

The above Figure 2.2 illustrates an AODV route lookup session. Node A wants to initiate traffic to node J for which it has no route. A transmit of a RREQ has been done, which is flooded to all nodes in the

network. When this request is forwarded to J from H, J generates a RREP. This RREP is then unicasted back to A using the cached entries in nodes H, G and D.

AODV builds routes using a route request/route reply query cycle. When a source node desires a route to a destination for which it does not already have a route, it broadcasts a route request (RREQ) packet across the network. Nodes receiving this packet update their information for the source node and set up backwards pointers to the source node in the route tables. In addition to the source node's IP address, current sequence number, and broadcast ID, the RREQ also contains the most recent sequence number for the destination of which the source node is aware. A node getting the RREQ may send a route reply (RREP) if it is either the destination or if it has a route to the destination with corresponding sequence number greater than or equal to that contained in the RREQ. If this is the case, it unicast a RREP back to the source. Otherwise, it rebroadcasts the RREQ. Nodes keep track of the RREQ's source IP address and broadcast ID. If they receive a RREQ which they have already processed, they discard the RREQ and do not forward it.

As the RREP propagates back to the source, nodes set up forward pointers to the destination. Once the source node receives the RREP, it may begin to forward data packets to the destination. If the source later receives a RREP containing a greater sequence number or contains the same sequence number with a smaller hop count, it may update its routing information for that destination and begin using the better route.

As long as the route remains active, it will continue to be maintained. A route is considered active as long as there are data packets periodically travelling from the source to the destination along that path. Once the source stops sending data packets, the links will time out and eventually be deleted from the intermediate node routing tables. If a link break occurs while the route is active, the node upstream of the break propagates a route error (RERR) message to the source node to inform it of the now unreachable destinations. After receiving the RERR, if the source node still desires the route, it can reinitiate route discovery[2].

Multicast routes are set up in a similar manner. A node wishing to join a multicast group broadcasts a RREQ with the destination IP address set to that of the multicast group and with the 'J'(join) flag set to indicate that it would like to join the group. Any node receiving this RREQ that is a member of the multicast tree that has a fresh enough sequence number for the multicast group may send a RREP. As the RREPs propagate back to the source, the nodes forwarding the message set up pointers in their multicast route tables. As the source node receives the RREPs, it keeps track of the route with the freshest sequence number, and beyond that the smallest hop count to the next multicast group member. After the specified discovery period, the source nodes will unicast a Multicast Activation (MACT) message to its selected next hop. This message serves the purpose of activating the route. A node that does not receive this message that had set up a multicast route pointer will timeout and delete the pointer. If the node receiving the MACT was not already a part of the multicast tree, it will also have been keeping track of the best route from the RREPs it received. Hence it must also unicast a MACT to its next hop, and so on until a node that was previously a member of the multicast tree is reached. AODV maintains routes for as long as the route is active. This includes maintaining a multicast tree for the life of the multicast group. Because the network nodes are mobile, it is likely that many link breakages along a route will occur during the lifetime of that route.

The counting to infinity problem is avoided by AODV from the classical distance vector algorithm by using sequence numbers for every route. The counting to infinity problem is the situation where nodes update each other in a loop.

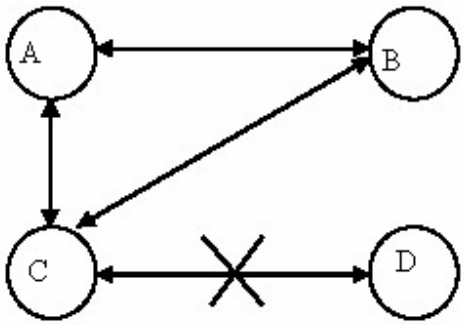


Figure 2.3 ―counting to infinity problem

Consider nodes A, B, C and D making up a MANET as illustrated in Figure2.3. A is not updated on the fact that its route to D via C is broken. This means that A has a registered route, with a metric of 2, to D. C has registered that the link to D is down, so once node B is updated on the link breakage between C and D, it will calculate the shortest path to D to be via A using a metric of 3.C receives information that B can reach D in 3 hops and updates its metric to 4 hops. A then registers

an update in hop-count for its route to D via C and updates the metric to 5. So they continue to increment the metric in a loop.

The way this is avoided in AODV, for the example described, is by B noticing that as route to D is old based on a sequence number. B will then discard the route and C will be the node with the most recent routing information by which B will update its routing table.

Characteristics of AODV

- Unicast, Broadcast, and Multicast communication.
- On-demand route establishment with small delay.
- Multicast trees connecting group members maintained for lifetime group.
- Link breakages in active routes efficiently repaired.
- All routes are loop-free through use of sequence numbers.
- Use of Sequence numbers to track accuracy of information.
- Only keeps track of next hop for a route instead of the entire route.
- Use of periodic HELLO messages to track neighbors.

Advantages and Disadvantages

The main advantage of AODV protocol is that routes are established on demand and destination sequence numbers are used to find the latest route to the destination. The connection setup delay is less. The HELLO messages supporting the routes maintenance are range-limited, so they do not cause unnecessary overhead in the network.

One of the disadvantages of this protocol is that intermediate nodes can lead to inconsistent routes if the source sequence number is very old and the intermediate nodes have a higher but not the latest destination sequence number, thereby having stale entries. Also multiple Route Reply packets in response to a single Route Request packet can lead to heavy control overhead. Another disadvantage of AODV is that the periodic beaconing leads to unnecessary bandwidth consumption.

2) Dynamic State Routing (DSR)

The DSR protocol requires each packet to carry the full address every hop in the route, from source to the destination. This means that the protocol will not be very effective in large networks, as the amount of overhead carried in the packet will continue to increase as the network diameter increases.

Therefore, in highly dynamic and large networks the overhead may consume most of the bandwidth. However, this protocol has a number of advantages over other routing protocols, and in small to moderately size networks (perhaps up to a few hundred nodes), this protocol performs better. An advantage of DSR is that nodes can store multiple routes in their route cache, which means that the source node can check its route cache for a valid route before initiating route discovery, and if a valid route is found there is no need for route discovery. This is very beneficial in network with low mobility, because the routes stored in the route cache will be valid for a longer period of time. Another advantage of DSR is that it does not require any periodic beaconing (or hello message exchanges), therefore nodes can enter sleep node to conserve their power. This also saves a considerable amount of bandwidth in the network. A full description of this protocol appears in later text.

Dynamic Source Routing (DSR) is a routing protocol for wireless mesh networks.[5] It is similar to AODV in that it establishes a route on-demand when a transmitting mobile node requests one. However, it uses source routing instead of relying on the routing table at each intermediate device.

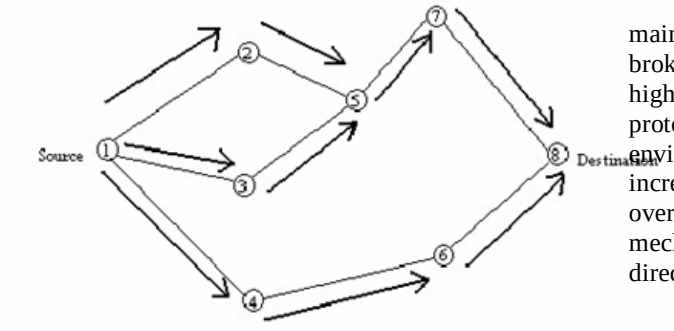
Dynamic source routing protocol (DSR) is an on-demand, source routing protocol, whereby all the routing information is maintained at mobile nodes. DSR allows the network to be completely self-organizing and self-configuring, without the need for any existing network infrastructure or administration. The protocol is composed of the two main mechanisms of "Route Discovery" and "Route Maintenance", which work together to allow nodes to discover and maintain routes to arbitrary destinations in the ad hoc network[5].

An optimum path for a communication between a source node and target node is determined by Route Discovery process. Route Maintenance ensures that the communication path remains optimum and loop-free according the change in network conditions, even if this requires altering the route during a transmission. Route Reply would only be generated if the message has reached the projected destination node, route record which is firstly contained in Route Request would be inserted into the Route Reply.

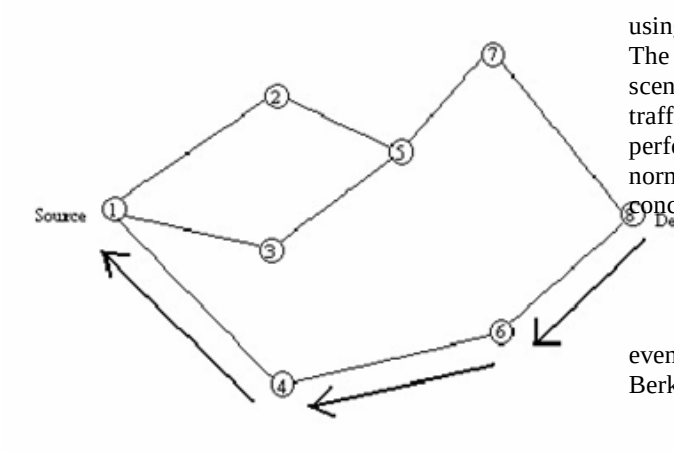
To return the Route Reply, the destination node must have a route to the source node. If the route is in the route cache of target node, the route would be used. Otherwise, the node will reverse the route based on the route record in the Route Reply message header -

symmetric links. In the event of fatal transmission, the Route Maintenance Phase is initiated whereby the Route Error packets are generated at a node. The incorrect hop will be detached from the node's route cache; all routes containing the hop are reduced at that point. Again, the Route Discovery Phase is initiated to determine the most viable route.

The major dissimilarity between this and the other on-demand routing protocols is that it is beacon-less and hence it does not have need of periodic hello packet transmissions, which are used by a node to inform its neighbors of its presence. The fundamental approach of this protocol during the route creation phase is to launch a route by flooding RouteRequest packets in the network. The destination node, on getting a RouteRequest packet, responds by transferring a RouteReply packet back to the source, which carries the route traversed by the RouteRequest packet received.



(a). Propagation of request (PREQ) packet



(b). Path taken by the Route Reply (RREP) packet

Figure 2.4 Creation of route in DSR

A destination node, after receiving the first RouteRequest packet, replies to the source node through the reverse path the RouteRequest packet had traversed. Nodes can also be trained about the neighboring routes traversed by data packets if operated in the promiscuous mode. This route cache is also used during the route construction phase. If an intermediary node receiving a RouteRequest has a route to the destination node in its route cache, then it replies to the source node by sending a RouteReply with the entire route information from the source node to the destination node.

Advantages and Disadvantages

DSR uses a reactive approach which eliminates the need to periodically flood the network with table update messages which are required in a table-driven approach. The intermediate nodes also utilize the route cache information efficiently to reduce the control overhead.

The disadvantage of DSR is that the route maintenance mechanism does not locally repair a broken down link[5]. The connection setup delay is higher than in table driven protocols. Even though the protocol performs well in static and low-mobility environments, the performance degrades rapidly with increasing mobility. Also, considerable routing overhead is involved due to the source-routing mechanism employed in DSR. This routing overhead is directly proportional to the path length.

3. Related Work

The performance differentials are analyzed using varying network load, mobility and network size. The internet connectivity may frequently create scenarios of multiple sources with constant bit rate traffic leading to common destination. In this paper, the performance of AODV and DSR are compared in normal and constrained scenarios for getting conclusions[7].

4. Simulation Environment

We simulated on Network Simulator[6], an event driven network simulator developed at UC Berkeley that simulates variety of IP networks.

To setup a simulation network, an OTCL script is written and to simulate it the script is executed which initiates an event scheduler and the network topology is setup using the network objects, controlling the traffic sources and the time to start and stop the transmitting of packets.

NAM and Xgraph are used for running the simulations. The biggest advantage of network animator (NAM) is that it provides a graphical user interface (GUI) for the different simulation environment according to the parameters specified by the user. The Xgraph utility generates the graphical output of the input data (or trace files).

To evaluate the performance of protocol in MANET, the protocol should be tested under realistic conditions such as – transmission range, data traffic, movement of mobile users etc. The simulations here use the Random Waypoint Mobility Model. It includes pause times between changes in destination and speed. The Random Waypoint model uses the concepts of epoch and pause making it a little bit more similar to realistic user mobility model.

Evaluations based on throughput and Packet delivery fraction comparison.

For all the simulations, the same movement model is used. The maximum speed of the nodes was set to 20m/s.
The number of nodes varied for 10 and 15.

The Simulation time was varied as 10s and 25s.

Simulation Scenario for Simulation time as 10s and 25s for 10 nodes

Parameter	Value
Number of nodes	**10**
Simulation Time	10 sec and 25 sec
Pause Time	5ms
Environment Size	800x800
Transmission Range	250 m
Traffic Size	Constant Bit Rate
Packet Size	512 bytes
Packet Rate	5 packets/s
Maximum Speed	20 m/s
Queue Length	50
Simulator	ns-2.34
Mobility Model	Random Waypoint
Antenna Type	Omni directional

Simulation Scenario for Simulation time as 10s and 25s for 15 nodes

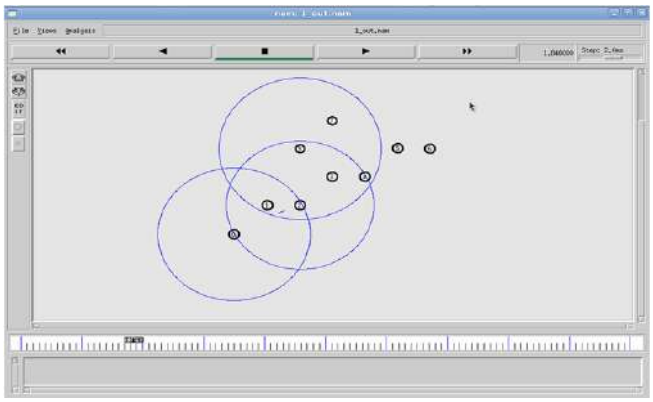
Parameter	Value
Number of nodes	**15**
Simulation Time	10 sec and 25sec
Pause Time	5ms
Environment Size	800x800
Transmission Range	250 m
Traffic Size	Constant Bit Rate
Packet Size	512 bytes
Packet Rate	5 packets/s
Maximum Speed	20 m/s
Queue Length	50
Simulator	ns-2.34
Mobility Model	Random Waypoint
Antenna Type	Omni directional

5. Simulation Results

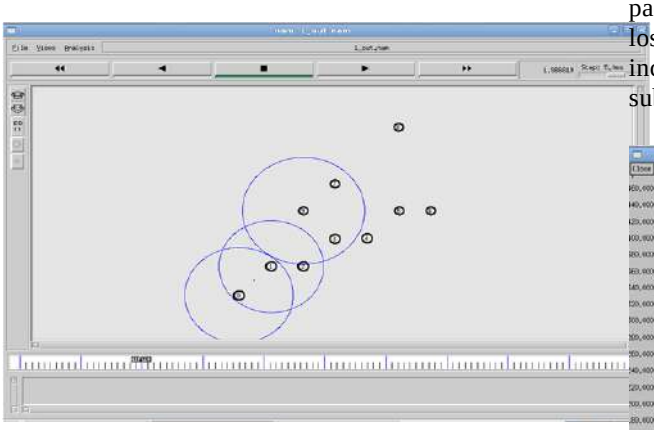
Experiments are conducted with CBR traffic sources sessions between different to common destination using AODV and DSR. The performance metric is Average Packet delivery rate. It was observed that AODV performs better than DSR in normal case. In another experiment four different CBR traffic sources started sessions with a common destination. The performance comparisons reflect that AODV suffers degradation of 30% whereas DSR suffers 10% compared to the normal situation (shown in Graphs)[7,8]. On comparing their performances, it was observed that DSR performs better than AODV under the constrained situation[7].

Simulation Graphs

1) A Screenshot of 10 nodes of AODV NAM – Network animator

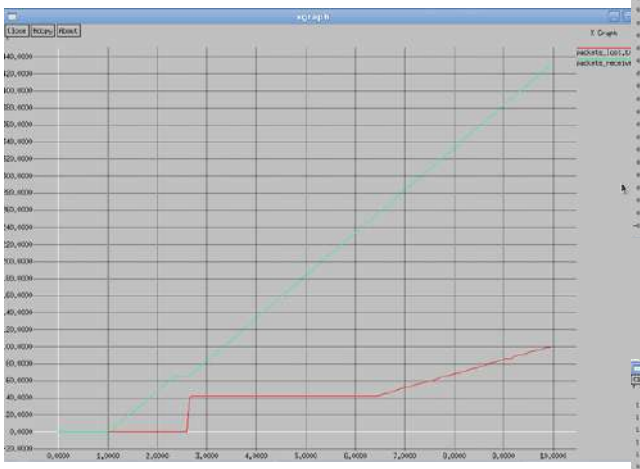


2) A Screenshot of 10 nodes of DSR NAM – Network animator



3) X Graph of 10 seconds simulation time of AODV with 10 nodes

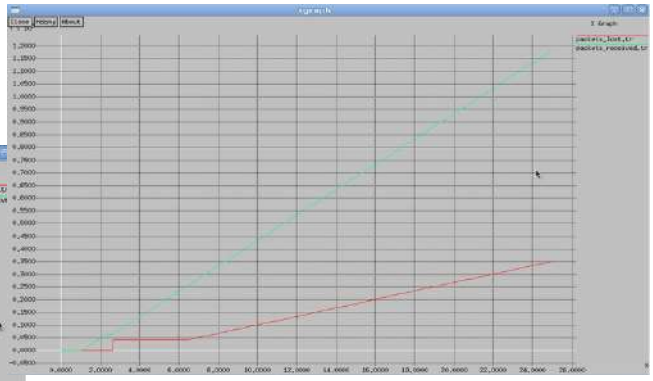
From fig, as the simulation start the packet received and packet loss is initially zero, because initially there is no CBR connection and nodes taking their right place. As the CBR connections establish between the nodes the number of packet received increases but no packet loss is there, it means all generated packets are being received by the nodes. But the packet loss increases substantially on the simulation time increases. Finally the packet received is more than the packet loss.



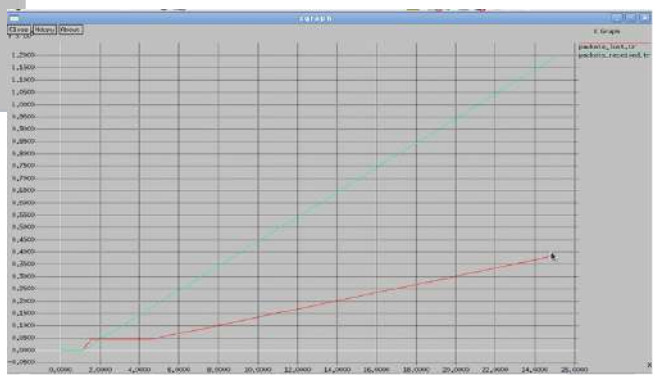
4) X Graph of and 10 seconds simulation time of DSR with 10 nodes

By the Figure we see that as the simulation start the packet received and packet loss is initially zero, because initially there is no CBR connection and nodes taking their right place. As the CBR connections establish the number of packet lost increases very much as compare to packet received. It shows that mostly generated packets are being dropped by the nodes. But the packet loss decreases substantially on the simulation time increases, and number of packet received increases substantially on the simulation time increases.

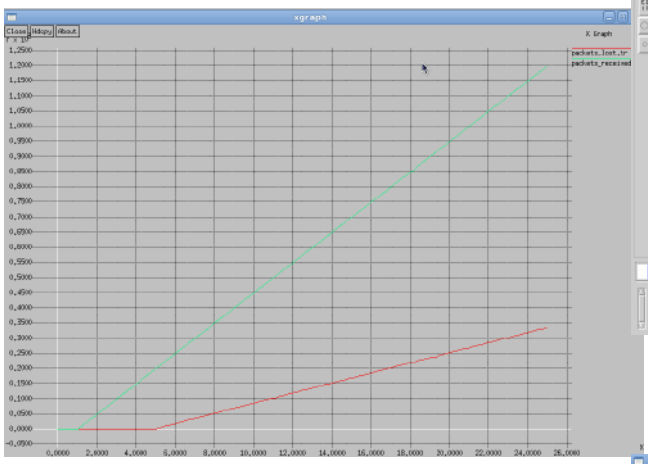
5) X Graph of 25 seconds simulation time of AODV with 10 nodes



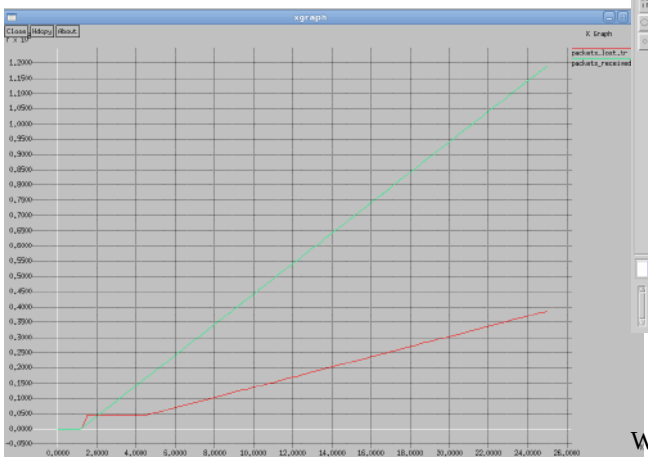
6) X Graph of 25 seconds simulation time of DSR with 10 nodes



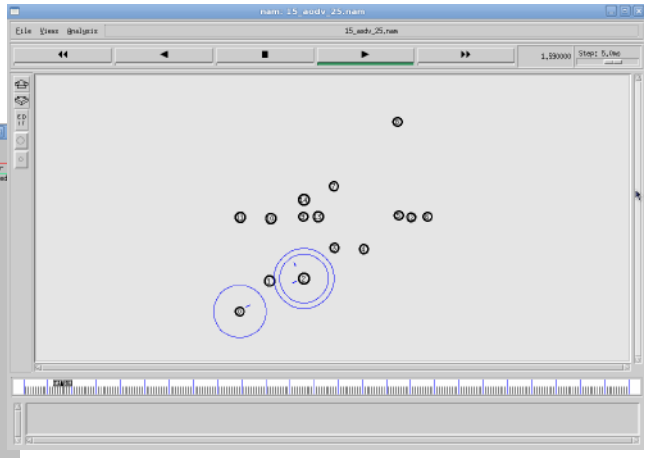
7) X Graph of 25 seconds simulation time of AODV with 15 nodes



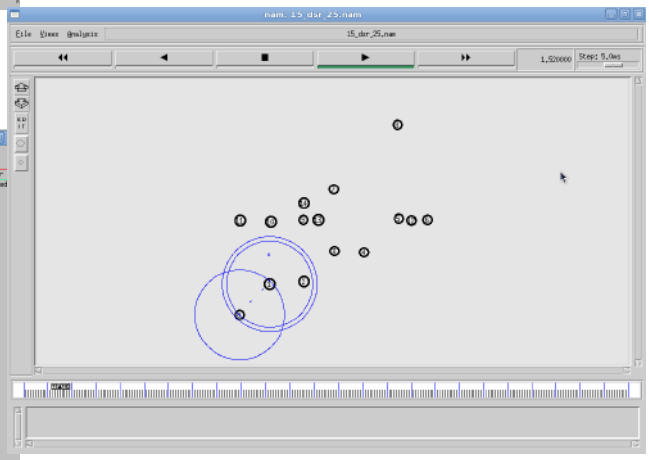
8) X Graph of 25 seconds simulation time of DSR with 15 nodes



9) A Screenshot of 15 nodes of AODV NAM – Network animator



10) A Screenshot of 15 nodes of AODV NAM – Network animator



6. CONCLUSIONS

We could observe that the performance of reactive routing protocol depends upon the scenario. In normal cases AODV performs better than DSR using various performance metrics. DSR works better than AODV in constrained situation of several CBR traffic sources leading to same destination in the mobile communicating nodes[7].

We could also see the same behavior of AODV & DSR in fact of packet receiving and packet loss. Initially there is no packet loss in AODV and a very high packet loss in DSR. But as the simulation time increases, the packet loss decreases and the packet receiving increases[7].

Moreover it can also be seen that AODV performs better with dynamic topology whereas DSR performs over AODV with static topologies.

ACKNOWLEDGMENT

No volume of words is enough to express my gratitude towards my guide, **Prof. Bhavika Gambhava**, Assist. Prof. , Department of Computer Engineering, Dharamsinh Desai University, who have been very concerned and direct me towards to research topic. They helped me to explore this vast topic in an organized manner and provided me with all the ideas on how to work towards to the research-oriented venture.

I am also thankful to **Prof. C. K. Bhansdadia,** Professor and Head of Department of Computer Science and Engineering, for the motivation and inspiration that triggered me for the research work.

I would like to say special thanks to **Dr. N. M. Bhatt,** Director, Gandhinagar Institute of Technology, who provide me the opportunity for presenting a paper in **GIT JET.**

I am thankful to **Prof. Darshana Mistry**, Co-ordinator of Computer Department at **Gandhinagar Institute of Technolgy**, for the motivation and inspiration for paper presentation.

I would like to thanks to **Ms. Happy Patel**, Lectuer, Gandhinagar Institute of Technology, for providing me required help for the completion of work.

Most importantly, I would like to thank my parents and the almighty for showing me the right direction out of the blue, to help me stay calm in the oddest of the times and keep moving even at times when there was no hope.

References

[1] C. E. Perkins. "Ad hoc Networking, Addison-Wesley",2001

[2] C. K. Toh, "Ad Hoc Mobile Wireless Networks: Protocols and Systems", Prentice Hall PTR,2002

[3] C. Perkins, E. Belding Royer and S. Das "Ad hoc on demand distance vector (AODV) routing." http://www.ietf.org/rfc/rfc3561txt, July 2003 IETF RFC 3561.

[4] C. Perkins, “Ad hoc on demand distance vector (AODV)routing” citeseer.nj.nec.com/article/perkins99ad.html, Nov. 1997.

[5] D. Johnson, D. Maltz and Y. Chun Hu. "The dynamic source routing protocol for mobile ad hoc networks." http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-10.txt, July 2004.IETF Internet Draft.

[6] “Network Simulator” http://www.isi.edu/nsnam/ns.

[7] R. Misra and C. R. Mandal, "Performance Comparison of AODV/DSR On-demand Routing Protocols for Ad Hoc Networks", in Constrained Situation, Proceedings of 7th IEEE International Conference on Personal Wireless Communications (ICPWC-2005).

[8] S. R. Das, C. E. Perkins, and E. M. Royer, “Performance comparison of two on-demand routing protocols for ad hoc networks, in Proc. INFOCOM 2000, pp3-12.

Margil P. Shah (M-10) become member of IEEE in 2010. He was born in Kalol 13- Dec-1984. He completed his B.E. computer engineering, L.C. Institute of Technology, Bhandu, Gujarat, India in 2007. **Bhumi S. Jani** was born in Bhavnagar 04-Dec-1986. She completed her B.E. Information of technology, S.V. Patel of Engineering, Visnagar, Gujarat, India in 2008.

He was working as Lecturer in Kalol Institute of Technology and Research Centre. He was Visiting Faculty in NIRMA Diploma Studies. He is working as Lecturer in computer department at Gandhinagar Institute of Technology from 2010. She is working as Lecturer in Kalol Institute of Technology and Research Centre since last 2 years.

Oracle Real Application Clusters

Nisha A Patel

Information Technology Engineering Department, Gandhinagar Institute of Technology
Gandhinagar, India
nisha.patel@git.org.in

***Abstract*-In database computing, Oracle Real Application Clusters (RAC) — an option for the Oracle Database software produced by Oracle Corporation and introduced in 2001 with Oracle9i — provides software for clustering and high availability in Oracle database environments. Oracle Corporation includes RAC with the Standard Edition of Oracle Database (aka Baby RAC), but makes it an extra-charge option for the Enterprise Edition**.

I. INTRODUCTION

ORACLE Real Application Clusters (RAC) allows Oracle Database to run any packaged or custom application, unchanged across a set of clustered servers. This provides the highest levels of availability and the most flexible scalability. If a clustered server fails, Oracle continues running on the remaining servers. When you need more processing power, simply add another server without taking users offline. To keep costs low, even the highest-end systems can be built out of standardized, commodity parts.

Oracle Real Application Clusters provides a foundation for Oracle's Enterprise Grid Computing Architecture. Oracle RAC technology enables a low-cost hardware platform to deliver the highest quality of service that rivals and exceeds the levels of availability and scalability achieved by the most expensive, mainframe SMP computers. By dramatically reducing administration
Costs and providing new levels of administration flexibility, Oracle is enabling the enterprise grid environment. Oracle RAC 11g Release 2 enables customers to build a dynamic grid infrastructure for their enterprise.

This paper provides a technical overview of Oracle Real Application Clusters 11g with the emphasis on the features and functionality that can be implemented to provide the highest availability and scalability for enterprise applications.

Oracle Real Application Clusters (RAC) is a fantastic achievement in the annals of relational database technology. But like any complex new technology, the adoption curve and skill set level have lagged slightly behind.

This had led to an interesting phenomenon that we refer to as "RAC in the Box Syndrome." Often people initially perform a proof-of-concept project with RAC and expect the cluster to perform "out of the box" as promised. Even though they have little experience optimizing RAC databases and focus so much on Oracle, they fail to optimize all subsystems in the "big picture." Disappointed with the initial test, many people prematurely abandon RAC and fall back to using symmetric multi-processing (SMP) boxes—the evil they already know and are comfortable using.

Oracle Real Application Clusters (RAC) is an option to the award-winning Oracle Database Enterprise Edition. Oracle RAC is a cluster database with a shared cache architecture that overcomes the limitations of traditional shared-nothing and shared-disk approaches to provide highly scalable and available database solutions for all your business applications. Oracle RAC is a key component of Oracle's enterprise grid architecture. Oracle RAC support is included in the Oracle Database Standard Edition for higher levels of system uptime.

Oracle RAC One Node is a new option available with Oracle Database 11*g* Release 2. Oracle RAC One Node is a single instance of an Oracle RAC enabled database running on one node in a cluster. This option adds to the flexibility that Oracle offers for database consolidation while reducing management overhead by providing a standard deployment for Oracle Databases in the enterprise.

II. RAC ARCHITECTURE

A. Overview of Oracle Real Application Clusters

A cluster comprises multiple interconnected computers or servers that appear as if they are one server to end users and applications. Oracle Real Application Clusters (Oracle RAC) enables you to cluster Oracle databases. Oracle RAC uses Oracle Cluster ware for the infrastructure to bind multiple servers so they operate as a single system.

Oracle Cluster ware is a portable cluster management solution that is integrated with the Oracle database. Oracle Cluster ware is also a required component for using Oracle RAC. In addition, Oracle Cluster ware enables both single-instance Oracle databases and Oracle RAC databases to use the Oracle high-availability infrastructure. Oracle Cluster ware enables you to create a clustered pool of storage to be used by any combination of single-instance and Oracle RAC databases.

Oracle Clusterware is the only Clusterware that you need for most platforms on which Oracle RAC operates. You can also use Clusterware from other vendors if the Clusterware is certified for Oracle RAC.

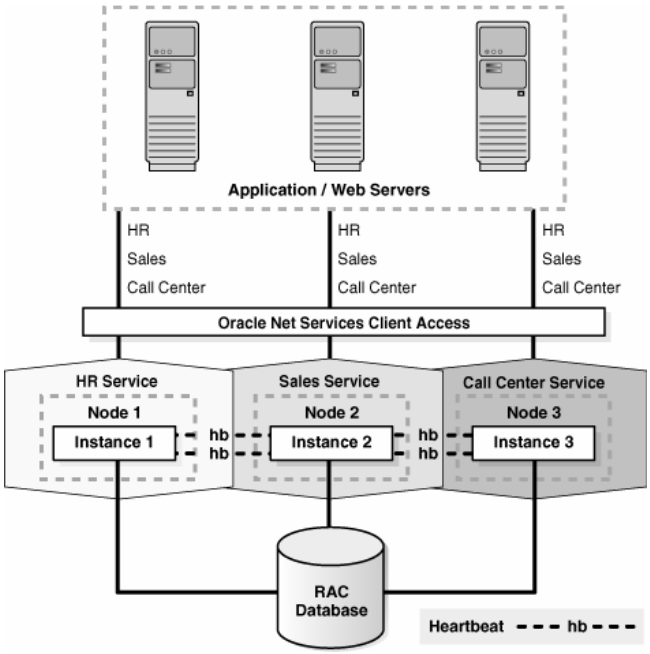


Fig. 1. *Overview of Oracle Clusterware for Oracle Real Application Clusters.*

B. Overview of Oracle Clusterware for Oracle Real Application Clusters

Oracle Clusterware provides a complete, integrated clusterware management solution on all Oracle Database platforms. This clusterware functionality provides all of the features required to manage your cluster database including node membership, group services, global resource management, and high availability functions. You can install Oracle Clusterware independently or as a prerequisite to the Oracle RAC installation process. Oracle database features such as services use the underlying Oracle Clusterware mechanisms to provide their capabilities. Oracle also continues to support select third-party clusterware products on specified platforms.

Oracle Clusterware is designed for, and tightly integrated with, Oracle RAC. When you create an Oracle RAC database using any of the management tools, the database is registered with and managed by Oracle Clusterware, along with the other Oracle processes such as Virtual Internet Protocol (VIP) address, Global Services Daemon (GSD), the Oracle Notification Service (ONS), and the Oracle Net listeners. These resources are automatically started when Oracle Clusterware starts the node and automatically restarted if they fail. The Oracle Clusterware daemons run on each node.

You can use Oracle Clusterware to manage high-availability operations in a cluster. Anything that Oracle Clusterware manages is known as a CRS resource, which could be a database, an instance, a service, a listener, a VIP address, an application process, and so on. Oracle Clusterware manages CRS resources based on the resource's configuration information that is stored in the Oracle Cluster Registry (OCR). You can use SRVCTL commands to administer other node resources. Oracle Clusterware stores the information that describes the configuration of these components in the OCR that you can administer as described in the Oracle Clusterware Administration and Deployment Guide.

C. Overview of Oracle Real Application Clusters Architecture and Processing

At a minimum, Oracle RAC requires a cluster software infrastructure that can provide concurrent access to the same storage and the same set of data files from all nodes in the cluster, a communications protocol for enabling interprocess communication (IPC) across the nodes in the cluster, enable multiple database instances to process data as if the data resided on a logically combined, single cache, and a mechanism for monitoring and communicating the status of the nodes in the cluster.

D. Overview of Oracle Real Application Clusters Installation and Database Creation

Install the grid infrastructure (which includes Oracle Clusterware and ASM) and Oracle Database software using Oracle Universal Installer, and create your database with Database Configuration Assistant (DBCA). This ensures that your Oracle RAC environment has the optimal network configuration, database structure, and parameter settings for the environment that you selected. As a database administrator, after installation your tasks are to administer your Oracle RAC environment at three levels:

a) Instance Administration
b) Database Administration
c) Cluster Administration (database administrators may or may not have privileges for cluster administration

Once you install the grid infrastructure and Oracle Clusterware is operational, run Oracle Universal Installer to install the Oracle database software with Oracle RAC components.

During the installation, Oracle Universal Installer runs DBCA to create your Oracle RAC database according to the options that you select. DBCA also runs the Net Configuration Assistant (NETCA) to configure the network for your Oracle RAC environment.

Oracle RAC software is distributed as part of the Oracle Database installation media. By default, the standard Oracle Database software installation process installs the Oracle RAC option when it recognizes that you are performing the installation on a cluster. The OUI installs Oracle RAC into a directory structure, which can be referred to as the Oracle home, which is separate from other Oracle software running on the system. Because OUI is cluster aware, it installs Oracle RAC software on all of the nodes that you defined to be part of the cluster.

Oracle recommends that you select ASM during the installation to simplify storage management; ASM automatically manages the storage of all database files within disk groups. If you are using Oracle Database Standard Edition, then you must use ASM to store all of the database

files. You can also configure services during installation, depending on your processing requirements.

By default, Oracle Database creates one service for your environment and the service is for the database. (The default database service is typically identified using the combination of the DB_NAME and DB_DOMAIN initialization parameters: db_name.db_domain.) The default service is available on all instances in an Oracle RAC environment, unless the database is in restricted mode. Oracle recommends that you create at least one service, in addition to the default database server, and use that service so that you have full control to manage your workload.

III. RAC CONFIGURATION AND IMPLEMENTATION

When you move to a RAC configuration, multiple instances on separate servers access the same database. In the diagram, database DB1 is deployed using a RAC configuration. The Instances are named DB11 and DB12.

The online redo logs, data files, control files and parameter files must reside on shared storage. Shared storage that is written to cooperatively from multiple servers at the same time

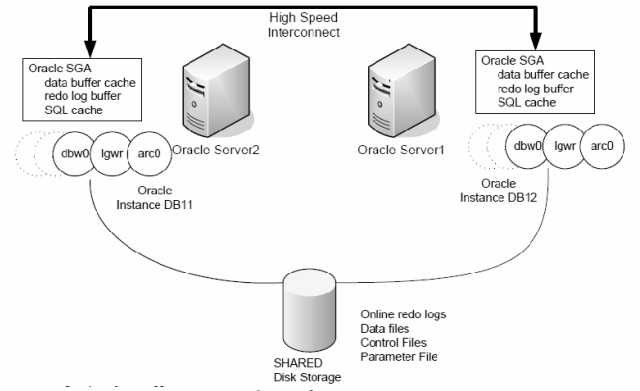


Fig. 2. *instance RAC Oracle configuration*.

is at the heart of RAC. Storage by default is not designed to be shared in this way.

In order ensure that writes are performed in the right sequence, the multiple instances must co-operate. For example, what happens if one instance needs to process a table row and that row is present in a block in the other cache, and perhaps hasn't been written to disk yet? That's where the high speed interconnect comes into play.

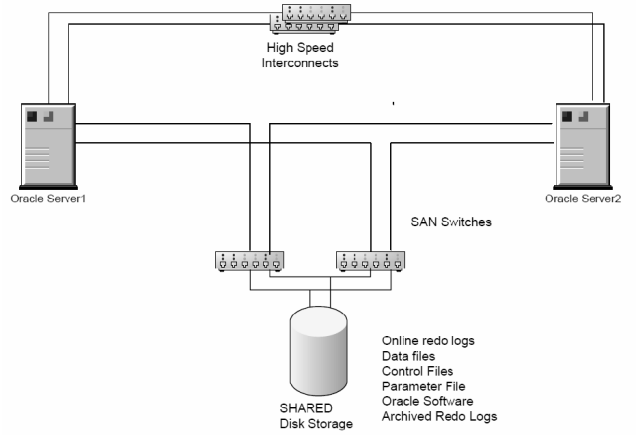


Fig. 3. instance RAC Oracle implementation.

Oracle manages this logical view through a process known as Global Cache Fusion (GCF). GCF ensures that data blocks are available where and when required. It's also at the heart of some performance benefits from RAC configurations, especially for applications that are more read than write intensive. For example, it's typically faster to fetch a block from the cache in the other instance across the interconnect than perform a physical read from the local instance. The downside is also true: for situations where there is need to update a particular block simultaneously on both nodes, RAC introduces the potential for additional overhead due to the need to synchronize the transfer of the block.

This diagram omits two essential components of the Oracle configuration – the location of the archived redo logs and the Oracle software itself. You need to decide where to locate them and that's not necessarily an easy decision, because each solution has costs and benefits.

As a high-availability/high-performance solution, your RAC deployment needs to have HA facilities and high-throughput capability built in through-out the stack. The following diagram shows an example of the hardware duplication required for a two instance cluster.

IV. CHALLENGES IN ADOPTING RAC TECHNOLOGY

One of the main challenges with any new technology, including Oracle RAC, is deciding at which point in the life-cycle to adopt it, if at all. Typically, any announcement of new technology by the mainstream DBMS vendors is accompanied by an enormous amount of fanfare. The downside is that new technologies tend to have gaps, anomalies, and deficiencies when deployed in the real world. Some of these gaps were known at the time of original release, and perhaps didn't get the publicity they should have done, and others have emerged over time.

This section sets out to provide some background on RAC configurations, goals, and technologies in the real world – along with a review of how RAC has changed since the original release (and why), and to what extent it meets its original goals in the current release, 10g.

A. Business Drivers for RAC Deployment

Typical studies include one or more of the following potential benefits for deployment of RAC:

a) Significant price/performance benefits Our experience is that this requires an Intel based solution in order to leverage the performance and prices gains from lower cost commodity servers from vendors such as HP, Dell, and IBM, to offset increased license costs from the RAC option.

b) Lower Total Cost of Ownership Many deployments and case studies existing demonstrating lower TCO when deploying RAC, however with the incremental effort and

technology required for new adoptions we believe the correct approach here is to develop a technology strategy based around Availability, Scalability, Performance and Cost and then asses whether RAC actually delivers against those requirements for key services relating to DBMS provisioning. There are many incremental costs that appear to be lost in translation in the TCO studies we have reviewed.

c) More Efficient System Utilization More efficient system utilization is promised by grid-based architectures. However, our experiences are that this goal represents more of future benefit than reality at the moment. To demonstrate the gap between theory and practice, a following section shows the basic steps in adding a node to cluster in RAC today – it's far from automatic and instantaneous.

d) Better Scalability to Handle Growth RAC does have the potential to scale by adding nodes to a cluster. However, given that most clusters we see today contain 2-nodes, we view this as more of a future benefit than reality. Also keep in mind that some vendors with a history of product deployment in business critical applications place relatively low limits on the number of RAC nodes their products support.

V. BENEFITS OF ORACLE REAL APPLICATION CLUSTERS

A. High Availability

Oracle Real Application Clusters 11g provides the infrastructure for datacentre high availability. It is also an integral component of Oracle's High Availability Architecture, which provides best practices to provide the highest availability data management solution. Oracle Real Application Clusters provides protection against the main characteristics of high availability solutions.

B. Reliability

Oracle Database is known for its reliability. Oracle Real Application Clusters takes this a step further by removing the database server as a single point of failure. If an instance fails, the remaining instances in the cluster are open and active. Oracle Clusterware monitors all Oracle processes and immediately restarts any failed component.

C. Recoverability

Oracle Database includes many features that make it easy to recover from all types of failures. If an Instance fails in an Oracle RAC database, it is recognized by another instance in the cluster and recovery automatically takes place. Fast Application Notification, Fast Connection Failover and Transparent Application Failover make it easy for applications to mask component failures from the user.

D. Error Detection

Oracle Clusterware automatically monitors Oracle RAC databases as well as other Oracle processes (ASM, listener, etc) and provides fast detection of problems in the environment. Also it automatically recovers from failures often before anyone has noticed a failure has occurred. Fast Application Notification provides the ability for applications to receive immediate notification of cluster component failures and mask the failure from the user by resubmitting the transaction to a surviving node in the cluster.

E. Continuous Operations

Oracle Real Application Clusters provides continuous service for both planned and unplanned outages. If a node (or instance) fails, the database remains open and the application is able to access data. Most database maintenance operations can be completed without down time and are transparent to the user. Many other maintenance tasks can be done in a rolling fashion so application downtime is minimized or removed. Fast Application Notification and Fast Connection Failover assist applications in meeting service levels and masking component failures in the cluster.

F. Scalability

Oracle Real Application Clusters provides unique technology for scaling applications. Traditionally, when the database server ran out of capacity, it was replaced with a new larger server. As servers grow in capacity, they are more expensive. For databases using Oracle RAC, there are alternatives for increasing the capacity. Applications that have traditionally run on large SMP servers can be migrated to run on clusters of small servers. Alternatively, you can maintain the investment in the current hardware and add a new server to the cluster (or to create a cluster) to increase the capacity. Adding servers to a cluster with Oracle Clusterware and Oracle RAC does not require an outage and as soon as the new instance is started, the application can take advantage of the extra capacity. All servers in the cluster must run the same operating system and same version of Oracle but they do not have to be exactly the same capacity. Customers today run clusters that fit their needs whether they are clusters of servers where each server is a 2 CPU commodity server to clusters where the servers have 32 or 64 CPUs in each server.

V. RAC OBJECTIONS

A. RAC is too complicated

Older versions of RAC required third-party Clusterware except on Linux–In 10g, the CRS provides a unified Clusterware–User can still choose third-party Clusterware based on their particular requirements.

RAC is fairly complicated–Adds moving parts–Applications can behave differently under RAC–Adds skillsetrequirements.

Automation & standardization can help–Make adding and removing nodes seamless–Transparently move processing capacity around between clusters–Create standardized builds that are easily deployable.

B. RAC doesn't scale

10gR2 RAC on Linux has been demonstrated in a lab to scale linearly to more than 10 nodes.

One of GridApp's customers using 10gR1 RAC has scaled linearly with a production transactional system up to 9 nodes.

Primarily a leftover objection from Oracle 9i RAC.

C. RAC is too expensive.

D. Commodity hardware isn't reliable enough.

VII. NEED OF RAC

You probably need RAC because:

a) Commodity processing is cheap
b) Downtime is expensive
c) Incremental scalability is a huge win
d) This is where the world is heading

VIII. EXISTING SYSTEM OF RAC

Oracle Parallel Server (1996-2001)—The OPS architecture allowed for several Oracle instances to share a common set of database files. In case of instance failure, the surviving instances could take over processing. There was a significant performance issue with OPS because shared RAM blocks had to be "pinged" between instances, imposing an additional processing burden on the cluster.

Real Application Clusters (2001–present)—The RAC architecture allows many instances to share a single database, but it avoids the overhead of RAM block pinging. RAC has also been enhanced to work with Oracle's Transparent Application Failover (TAF) to automatically restart any connections when an instance fails.

To overcome the IDLM problem in OPS, Oracle overhauled the architecture of the OPS product and reintroduced it under a new name: RAC. RAC employs a new technology called Cache Fusion, whereby the data block buffers of all instances within the parallel server configuration reside in a single shared RAM memory region. By having all data blocks instantly available to all database instances, the problem of IDLM pinging is overcome, allowing the systems to run faster and with greater reliability than with OPS.

IX. CONCLUSION

ORACLE Real Application Clusters has been designed for high availability and scalability. By providing protection from hardware and software failures, Oracle Real Application Clusters provides systems availability ensuring continuous data access. Its scale out and scale up features offer a platform, which can grow in any direction allowing enterprises to grow their businesses. Existing applications as well as newly developed applications benefit from the transparency Oracle Real Application Clusters provides. Application development as well as administration and change management thus become much easier allowing reduction in total cost of ownership. Oracle Real Application Clusters is unique to the market with its offering and capabilities. Oracle RAC is used by thousands of customers worldwide in all industries in mission critical and many other application environments.

RAC offers huge advantages over traditional single-instance environments However, RAC is not the seamless transition Oracle sometimes claims Plan, document, test, and reap the rewards

REFERENCES

[1] http://www.oracle.com/us/products/database/options/real-application-clusters/index.htm
[2] http://www.oraclepoint.com/topic_455
[3] www.dba-oracle.com/real...rac.../advantages_of_rac.html
[4] http://en.wikipedia.org/wiki/Oracle_RAC
[5] http://www.nyoug.org/Presentations/2006/September_NYC_Metro_Meeting/200609Zito_You%20Probably%20DO%20Need%20RAC.pdf
[6] http://www.oracle.com/us/corporate/analystreports/industries/059641.pdf
[7] http://www.oracle.com/technetwork/database/.../index.html
[8] http://en.wikipedia.org/wiki/Oracle_RAC
[9] http://www.oracle.com/us/products/database/options/real-application-clusters/index.htm
[10] http://www.oracle.com/us/products/database/options/real-application-clusters/index.htm
[11] http://www.oraclepoint.com/topic_455
[12] *http://www.dba-**oracle**.com/real...**rac**.../advantages_of_**rac**.html*
[13] http://en.wikipedia.org/wiki/Oracle_RAC
http://www.nyoug.org/Presentations/2006/September_NYC_Metro_Meeting/200609Zito_You%20Probably%20DO%20Need%20RAC.pdf
[14] http://www.**oracle**.com/technetwork/database/.../index.html
[15] http://en.wikipedia.org/wiki/**Oracle_RAC**

Software Selection Criteria and Numerical Technique: NPV (Net Present Value)

Svapnil M Vakharia

Information Technology Engineering Department, Gandhinagar Institute of Technology
Gandhinagar, India
svapnil.vakharia@git.org.in

***Abstract*— This paper presents a broad view of software selection methodology and criteria for software selection techniques. In this paper I describe a software selection numerical technique NPV (net present value) along with the NPV calculation. Although Software selection is a critical area of research in the information systems field, relatively little such research has been published in leading IS journals. To point toward increased research opportunities for software selection work, this paper presents a framework that illustrates the broad range of criteria within the software selection area.**

***Index Terms*--Software selection, project selection model criteria, NPV, Project Evaluation factor**

I. INTRODUCTION

Project selection is the process of evaluating individual projects or groups of projects, and then choosing to implement some set of them so that the objectives of the parent organization will be achieved. This same systematic process can be applied to any area of the organization's business in which choices must be made between competing alternatives. Project selection is only one of many decisions associated with project management.

II PROJECT SELECTION MEDEL CRITERIA

A. Realism

The model should reflect the reality of the manager's decision situation, including the multiple objectives of both the firm and its managers. Without a common measurement system, direct comparison of different projects is impossible. For example, Project A may strengthen a firm's market share by extending its facilities, and Project B might improve its competitive position by strengthening its technical staff. Other things being equal, which is better? The model should take into account the realities of the firm's limitations on facilities, capital, personnel, and so forth. The model should also include factors that reflect project risks, including the technical risks of performance, cost, and time as well as the market risks of customer rejection and other implementation risks.

B Capability

The model should be sophisticated enough to deal with multiple time periods, simulate various situations both internal and external to the project (e.g., strikes, interest rate changes), and optimize the decision. An optimizing model will make the comparisons that management deems important, consider major risks and constraints on the projects, and then select the best overall project or set of projects.

C. Flexibility

The model should give valid results within the range of conditions that the firm might experience. It should have the ability to be easily modified, or to be self-adjusting in response to changes in the firm's environment; for example, tax laws change, new technological advancements alter risk levels, and, above all, the organization's goals change.

D. Ease of use

The model should be reasonably convenient, not take a long time to execute, and be easy to use and understand. It should not require special interpretation, data that are difficult to acquire, excessive personnel, or unavailable equipment. The model's variables should also relate one-to-one with those real-world parameters the managers believe significant to the project. Finally, it should be easy to simulate the expected outcomes associated with investments in different project portfolios.

E. Cost

Data-gathering and modeling costs should be low relative to the cost of the project and must surely be less than the potential benefits of the project. All costs should be considered, including the costs of data management and of running the model.

F. Easy computerization

It should be easy and convenient to gather and store the information in a computer database, and to manipulate data in the model through use of a widely available, standard computer package. we first examine fundamental types of project selection models and the characteristics that make any model more or less acceptable. Next we consider the limitations, strengths, and weaknesses of project selection models, including

some suggestions of factors to consider when making a decision about which, if any, of the project selection models to use. We then discuss the problem of selecting projects when high levels of uncertainty about outcomes, costs, schedules, or technology are present, as well as some ways of managing the risks associated with the uncertainties. Finally, we comment on some special aspects of the information base required for project selection.

III. **P**ROJECT **E**VALUTION **F**ACTORS

A. Production Factors

1. Time until ready to install
2. Length of disruption during installation
3. Learning curve—time until operating as desired
4. Effects on waste and rejects
5. Energy requirements
6. Facility and other equipment requirements
7. Safety of process
8. Other applications of technology
9. Change in cost to produce a unit output
10. Change in raw material usage
11. Availability of raw materials
12. Required development time and cost
13. Impact on current suppliers
14. Change in quality of output

B. Marketing Factors

1. Size of potential market for output
2. Probable market share of output
3. Time until market share is acquired
4. Impact on current product line
5. Consumer acceptance
6. Impact on consumer safety
7. Estimated life of output
8. Spin-off project possibilities

C. Financial Factors

1. Profitability, net present value of the investment
2. Impact on cash flows
3. Payout period
4. Cash requirements
5. Time until break-even
6. Size of investment required
7. Impact on seasonal and cyclical fluctuations

D. Personnel Factors

1. Training requirements
2. Labor skill requirements
3. Availability of required labor skills
4. Level of resistance from current work force
5. Change in size of labor force
6. Inter- and intra-group communication requirements
7. Impact on working conditions

E. Administrative and Miscellaneous Factors

1. Meet government safety standards
2. Meet government environmental standards
3. Impact on information system
4. Reaction of stockholders and securities markets
5. Patent and trade secret protection
6. Impact on image with customers, suppliers, and competitors
7. Degree to which we understand new technology
8. Managerial capacity to direct and control new process

Some factors in this list have a one-time impact and some recur. Some are difficult to estimate and may be subject to considerable error. For these, it is helpful to identify a range of uncertainty. In addition, the factors may occur at different times. And some factors may have thresholds, critical values above or below which we might wish to reject the project. Clearly, no single project decision need include all these factors. Moreover, not only is the list incomplete, it also contains redundant items. Perhaps more important, the factors are not at the same level of generality: profitability and impact on organizational image both affect the overall organization, but impact on working conditions is more oriented to the production system. Nor are all elements of equal importance. Change in production cost is usually considered more important than impact on current suppliers. Shortly, we will consider the problem of generating an acceptable list of factors and measuring their relative importance. At that time we will discuss the creation of a Decision Support System (DSS) for project evaluation and selection.

IV **I**NTRODUCTION TO **N**PV

.

In simple terms, NPV is the difference between an investment's market value and its costs (sum of the present values of the net annual benefits minus the initial dollar investment required). The "P" in NPV means that we will derive a single dollar value for the investment today even though the life of the project may span many years. We begin by estimating the future cash flows that the new software will produce. For purposes of these calculations, any cost saving, incremental revenues or other benefits to the organization as a result of deploying the software are considered "inflows" and are positive numbers. All costs are considered "outflows" and are negative numbers.

Positive NPV ≡ PV of outputs > PV of inputs

Positive NPV ≡ economic return > economic cost some causes of +NPV

V. **N**ET **P**RESENT **V**ALUE **C**ALCULATION

By definition, NPV = the sum of the present values of all of a project's cash flows, both negative (cash outflows) and positive (cash inflows). For the sake of simplicity, the project cash flows are estimated on an annual basis. The formula for calculating the NPV is:

$NPV_n = (PV_1 + PV_2 + ... + PV_n)$ – Initial Investment Cost

Where:

NPV_n = the Net Present Value of the project over n years

PV_1 through PV_n = the cash flows from each project year (positive for cash inflows, negative for cash outflows).

The formula for calculating the Present Value for a cash flow in a particular year is:

$PV_n = FV_n * PVF_{nd}$

Where:

PV_n = the Present Value of the cash flow from year n

FV_n = the known Future Value of the project cash flow in year n

PVF_{nd} = a Present Value Factor for the year (n) and the project discount rate (d)

Values of PVF have been calculated for various combinations of n and d and are organized on "Present Value Tables", where they can be looked up easily (a version has been included at the end of this instruction guide).

Before doing the NPV calculation for a project, you will need the following information:

1) The initial investment cost

2) The future cash inflows or outflows (FV) expected to occur in each subsequent year of the project. Sometimes the future cash flows will be the same every year, and sometimes they will be irregular. Sometimes they will be all cash inflows, and sometimes a mix of inflows and outflows. It will vary from project to project.
3) The discount rate (d) for the company or the project. Some companies use an average discount rate for the analysis of all projects. Other companies may prefer slightly different discount rates for different projects. The discount rate you use should be equal to the required rate of return for the investment project, and should take into account price inflation, project risk, and the real return that you require. At a minimum, this required rate of return should cover the cost of investment capital to the firm.
4) The number of years (n) over which you would like to estimate project profitability.

Using a "Present Values Table" to determine Present Value Factors

A Present Value Table will allow you to look up Present Value Factors (PVF) for various combinations of n (project year) and d (project discount rate). As an example, look at a CP investment with the following parameters:

1. Initial investment: US$150,000
2. Future savings (FV):
   Year 1— US$45,000
   Year 2— US$45,000
   Year 3— US$77,000
3. Discount rate (d): 10%
4. Number of years (n): 3

Using the Present Value table attached, look up the Present Value Factors (PVF) for a discount rate of 10% and for project years 1, 2, and 3.

Year 1 PVF: 0.9091
Year 2 PVF: 0.8264
Year 3 PVF: 0.7513

Using a calculator to calculate Present Value Factors

If you do not have a Present Value table available, you can calculate the necessary Present Value Factors yourself as follows:

Present Value factor (PVF) =1/ $(1+d)^n$

Where

d is the discount rate
n is the year number

Using this formula, and a discount rate of 10%, PVFs for years 1, 2, and 3 can be calculated as follows.

Year1PVF= **1/**$(1+0.1)^1$**= 0.9091**
Year2PVF= **1/**$(1+0.1)^2$**= 0.8264**
Year2PVF= **1/**$(1+0.1)^3$ **= 0.7513**

These are exactly the same PVFs that you looked up on the table previously. You would use these values to calculate NPV the same was as is illustrated above.

Using the PVFs shown above, the future cost savings for each year can be converted to their present value. These values are then added together to estimate the project's Net Present Value. The initial investment (which is already in present-day dollars) is subtracted from the sum. The result is the Net Present Value of the project.

Net Present Value Calculation Table

Year	Future Saving	Present Value Factor(10%)	Present Value
1	$45000	X 0.9091	$40,910
2	$45000	X 0.8264	$37,188
3	$77000	X 0.7513	$57,850
		Total	$135,948
	less: initial investment		- $150,000
	equals: **Net Present Value**		**-$14,052**

For this example, the NPV is calculated to be -$14,052 which means that the investment is not profitable within three years. A positive value for NPV would indicate that the investment is profitable within three years.

VI. CONCLUTION

Software Cost Estimation continues to be a weak link in software project management. The aim of this paper is to present a review of current cost estimation criteria to help both industry and academia in choosing the appropriate methods when preparing software cost estimates. The paper cover software selection technique called NPV (net present value) underlying aspects in preparing cost estimates. In terms of the estimation process, nearly all numerical models deviate from the classical view of the cost estimation process.

REFERENCES

[1] BROWN, S. L., and K. M. EISENHARDT. Competing on the Edge: Strategy as Structured Chaos.Boston: Harvard Business School Press, 1998.

[2] COOPER, R. G., S. J. EDGETT, and E. J. KLEINSCHMIDT.Portfolio Management for New Products. Reading, MA: Addison-Wesley, 1998.

[3] FRAME, J. D. The New Project Management: Toolsfor an Age of Rapid Change, Corporate Reengineering, and Other Business Realities. San Francisco: Jossey-Bass Publishers, 1994.

[4] GRAHAM, ROBERT J. and RANDALL L. ENGLUND.Creating an Environment for Successful Projects: The Quest to Manage Project Management. San Francisco: Jossey-Bass Publishers, 1997.

[5] KUMAR, V., et al. “To Terminate or Not an Ongoing R&D Project: A Managerial Dilemma.” IEEE Transactions on Engineering Management 279 (1996)

[6] MARTINO, J. R & D Project Selection. New York: Wiley, 1995.

[7] WESTNEY, R. E. Computerized Management of Multiple Small Projects. New York: Dekker, 1992.

Svapnil Vakharia(ISTE-LM’10). he became a life member ship of ISTE in 2010. His date of birth is 9st September 1984 in kapadvanj. he completed B.E.Infromation Technology Engineer from Saurastra .University(Gujarat) with 61.53% 2006 , M.IS from Bllarat University(sydney,Australia) with 68.13% in 2008.

He is working as**, Asst. Professor in Gandhinagar Institute of Technology college, Gandhinagar.**

Introduce a New Phase in System Development Life Cycle (SDLC)

Svapnil Vakharia

Assistant professor,
Department of Information Technology
Gandhinagar Institute of Technology,India
svapnil.vakharia@git.org.in

Abstract: - One of the greatest challenges facing IT project managers today is the seamless integration of project management methodology and SDLC methodology. Organizations that are developing custom software solution are faced with the difficult choice of picking the right software development life cycle (SDLC). This paper summarizes the steps an organization would have to go through in order to make the best possible choice. In this paper I also include a new Phase for software development life cycle. The SDLC model, I include a software selection phase along with each software selection step detail.

Key words: Software Development Life Cycle, Software Selection, SDLC Methodology, disadvantages of SDLC, Phases of SDLC

I. METHOLOGY OERVIEW

Before considering a framework for selecting a given SDLC methodology, one needs to understand the choices. In this paper I am focused on constructing SDLC Model and will seek to address the primary approach of the various methodologies in supporting software development initiatives. Systems development, systems engineering, integration, and Commercial off the Shelf (COTS) software package implementation initiatives are included within these life cycles as well, however the majority of our discussion will focus on software selection specifically. SDLC methodologies all focus on the common goal of defining the steps and processes from the beginning of a software project through its completion. Traditional engineering models have depicted these steps as Requirements Analysis, Design, Development, Test, and Implementation. Accordingly, the models herein presented add a step to the beginning (software Selection) captures the day-to-day ongoing activities necessary to sustain the system.

II. INTRODUCTION OF SYSTEM DEVELOPMENT LIFE CYCLE (SDLC)

A software cycle deals with various parts and phases from planning to testing and deploying software. All these activities are carried out in different ways, as per the needs. Each way is known as a Software Development Lifecycle Model (SDLC).A software life cycle model is either a descriptive or prescriptive characterization of how software is or should be developed. A descriptive model describes the history of how a particular software system was developed. Descriptive models may be used as the basis for understanding and improving software development processes or for building empirically grounded prescriptive models.

The multistep process that starts with the initiation, analysis, design, and implementation, and continues through the maintenance and disposal of the system, is called the System Development Life Cycle (SDLC). The system development life cycle is the overall process of developing, implementing, and retiring information systems through a multistep process from initiation, analysis, design, implementation, and maintenance to disposal. There are many different SDLC models and methodologies, but each generally consists of a series of defined steps or phases.

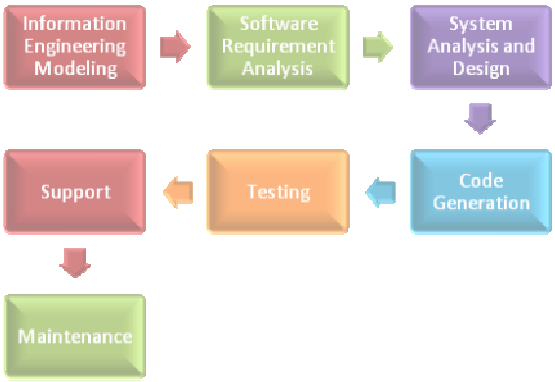
III. PHASES OF SYSTEM DEVELOPMENT LIFE CYCLE

(A)System/Information Engineering and Modeling

As software is always of a large system (or business), work begins by establishing requirements for all system elements and then allocating some subset of these requirements to software. This system view is essential when software must interface with other elements such as hardware, people and other resources. System is the basic and very critical requirement for the existence of software in any entity. So if the system is not in place, the system should be engineered and put in place. In some cases, to extract the maximum output, the system should be re-engineered and spruced up. Once the ideal system is engineered or tuned, the development team studies the software requirement for the system.

(B)Software Requirement Analysis

This is also known as feasibility study. In this phase, the development team visits the customer and studies their system. They investigate the need for possible software automation in the given system. By the end of the feasibility study, the team furnishes a document that holds the different specific recommendations for the candidate system. It also includes the personnel assignments, costs, project schedule, and target dates. The requirements gathering process is intensified and focused specially on software. To understand the nature of the program(s) to be built, the system engineer ("analyst") must understand the information domain for the software, as well as required function, behavior, performance and interfacing. The essential purpose of this phase is to find the need and to define the problem that needs to be solved.



(Phases in SDLC (Software Development Life Cycle))

(C) System Analysis and Design

In this phase, the software development process, the software's overall structure and its nuances are defined. In terms of the client/server technology, the number of tiers needed for the package architecture, the database design, the data structure design etc are all defined in this phase. A software development model is created. Analysis and Design are very crucial in the whole development cycle. Any fault in the design phase could be very expensive to solve in the later stage of the software development. Much care is taken during this phase. The logical system of the product is developed in this phase.

(D) Code Generation

The design must be translated into a machine-readable form. The code generation step performs this task. If the design is performed in a detailed manner, code generation can be accomplished without much complication. Programming tools like Compilers, Interpreters, and Debuggers are used to generate the code. Different high level programming languages like C, C++, Pascal, and Java are used for coding. With respect to the type of application, the right programming language is chosen.

(E) Testing

Once the code is generated, the software program testing begins. Different testing methodologies are available to unravel the bugs that were committed during the previous phases. Different testing tools and methodologies are already available. Some companies build their own testing tools that are tailor made for their own development operations.

(F)Maintenance

Software will definitely undergo change once it is delivered to the customer. There are many reasons for the change. Change could happen because of some unexpected input values into the system. In addition, the changes in the system could directly affect the software operations. The software should be developed to accommodate changes that could happen during the post implementation period.

IV **D**ISADVANTAGES OF **E**XISTING **S**DLC **M**ODELS

(A) Waterfall

(1)Lack of measurable progress within stages
(2)Cannot accommodate changing requirements
(3)Resistant to time and/or budget compression

(B) Incremental

(1)Demands increased management attention
(2) Can increase resource requirements
(3) No support for changing requirements

(c) Evolutionary

(1)Increases management complexity
(2) IOC only partially satisfies requirements and is not complete functionality
(3) Risk of not knowing when to end the project

(D) Spiral

(1) Increased management complexity
(2)Defers production capability to end of the SDLC
(3)Risk of not knowing when to end the project

(E) RAD (Rapid Application Development)

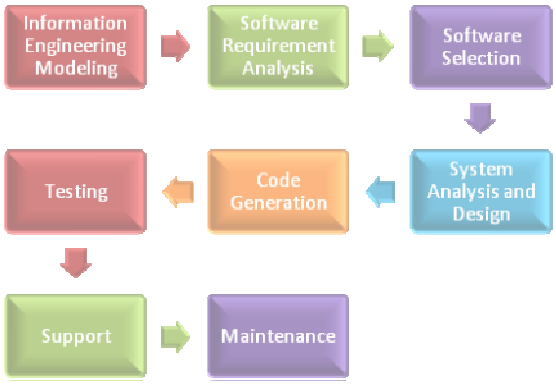
(1) Increases management complexity
(2) Drives costs forward in the SDLC
(3) Can increase resource requirements

(F) Extreme/Agile Development

(1) Not conducive to handling complex dependencies
(2)Creates Quality Assurance (QA) risks
(3)Increased risk of sustainability, maintainability, and extensibility

V **I**NTRODUCTION OF **S**OFTWARE **S**ELECTION **P**HASE IN **S**YSTEM **D**EVELOPMENT **L**IFE **C**YCLE

Introduce Software Selection Phase between software requirement analysis and System Analysis and Design:



(Selection Phase in SDLC (Software Development Life Cycle))

Software project Selection includes estimates of cost, product size, resources, staffing levels, schedules, and key milestones. The software estimation process discussed in the following subsections describes the steps for developing software estimates. Establishing this process early in the life-cycle will result in greater accuracy and credibility of estimates and a clearer understanding of the factors that influence software development costs.

Software Selection Steps:

Step 1 - Gather and Analyze Software Functional and Programmatic Requirements

The purpose of this step is to analyze and refine the software functional requirements and to identify technical and programmatic constraints and requirements that will be included in the software estimate. This enables the work elements of the project-specific WBS to be defined and software size and effort to be estimated. Analyze and refine the requirements as follows:

1. Analyze and refine the software functional requirements to the lowest level of detail possible. Clearly identify requirements that are not well understood in order to make appropriate risk adjustments. If an incremental development strategy is used, then the refinement will be based on the requirements that have been defined for each increment.
2. Analyze and refine a software physical architecture hierarchy based on the functional requirements. Define the architecture in terms of software segments to be developed. Decompose each segment to the lowest level function possible.
3. Analyze project and software plans to identify programmatic constraints and requirements including imposed budgets, schedules, margins, and make/buy decisions.

The outputs of this step are:
• Technical and programmatic constraints and requirements
• Assumptions made about the constraints and requirements
• Methods used to refine the software functional requirements
• Refined software functional requirements
• Software architecture hierarchy of segments and associated functions

Step 2 - Define the Work Elements and Procurements

The purpose of this step is to define the work elements and procurements for the software project that will be included in the software estimate.

1. Use the WBS to plan the work elements and procurements for the project that requires estimation. Then consult your project-specific WBS to find additional applicable work elements.
The work elements and procurements will typically fall into the following categories of a project-specific WBS:

- Software Management
- Software Systems Engineering
- Software Engineering
- Software Test Engineering
- Software Development Test Bed
- Software Development Environment
- Software System-level Test Support
- Assembly, Test, Launch Operations (ATLO) Support for flight projects
- SQA
- IV&V

These WBS categories include activities across the software life-cycle from requirements analysis through completion of system test. Note that software operations and support (including maintenance) is not in the scope of these estimates. Work elements such as SQA and IV&V are not often part of the software manager's budget, but are listed here to remind software managers that these services are being provided by the project.

2. Identify the attributes of the work elements that will drive the size and effort estimates in terms of heritage and risk. From this, derive an initial risk list.

- Anything that is new, such as code, language, or design method
- Low technology willingness levels
- Overly optimistic assumptions related to high custom elements
- Possible reuse
- Vendor-related risks associated with Commercial Off-The-Shelf (COTS) software
- Criticality of mission failure
- Software classification
- Use of development tools
- Concurrent development of hardware
- Number of interfaces between multiple development organizations
- Geographical distribution of multiple development organizations
- High complexity elements
- Skill and experience level of team
- Vague or incomplete requirements
- The outputs of this step include the following:

- Assumptions about the work elements and procurements
- List of procurements

Step 3 - Estimate Software Size

The purpose of this step is to estimate the size of the software product. Because formal cost estimation techniques require software size as an input size prediction is essential to effective effort estimation. However, size is often one of the most difficult and challenging inputs to obtain.
The most commonly used industry-wide measure of software size is the number of source lines of code (SLOC). Typically either physical lines or logical lines are used when counting SLOC. Comments and blanks should never be included in any count of lines of code. The physical SLOC measure is very simple to count because each line is terminated by the enter key or a hard line break. A logical statement is a single software instruction, having a defined beginning and ending independent of any relationship to the physical lines on which it is recorded or printed. Logical statements may encompass several physical lines and typically include executable statements, declarations, and compiler directives.
The outputs of this step are as follows:

- Assumptions made in order to estimate software size
- Methods used to estimate software size
- Software size estimates for each function and software heritage category in logical SLOC
- Total software size estimate in logical SLOC

Step 4 - Estimate Software Effort

The purpose of this step is to convert the software size estimates, from the previous step, to Software Development Effort. Software Development Effort covers software systems Engineering, test engineering, and software engineering work to develop the software from requirements analysis up through software I&T. If you have not completed a size estimate then Obtain effort data for analogous software tasks and functions, and apply the steps described under size estimation to derive the software development effort directly.
Size estimates are used to calculate effort in work-months (WM) for the Software Development work elements of the WBS. The Software Development work elements of the WBS include Software System Engineering, Software Engineering, and Software Test Engineering. The effort and cost for the other work elements are calculated in later steps using other methods.
Convert the size of each software function to Software Development Effort as follows:

1. *SW_Development_Effort = Size_Estimate / SW_Development_Productivity*

where,

• *SW_Development_Effort* is measured in WM.
• *SW_Development_Productivity* is measured in SLOC/WM.
• *Size_Estimate* is measured in logical SLOC.

2. Adjust the effort estimates of each software function for software heritage by multiplying the Software Development Effort by the effort multiplier according to Effort Adjustment Multipliers for Software Heritage One of the major causes of cost growth is optimistic software heritage assumptions.
Therefore, any reduction in effort based on software heritage should be viewed with caution. Nominally, projects have significant software design heritage, but require the writing of completely new code. If a project requires completely new design (not new technology) and new code to be developed, then it will require on average 20% more effort than the nominal case. If some code is being reused, effort can be decreased. New technology can increase effort by 50%-200%.
3. Sum the adjusted Software Development Effort of each function and software heritage category to arrive at the Total Software Development Effort.
The outputs of this step are as follows:

- Assumptions made in order to estimate Software Development Effort including heritage
- Methods used to estimate Software Development effort
- Software Development Effort of each function adjusted for heritage in work-months

Step 5 - Schedule the Effort

The purpose of this step is to determine the length of time needed to complete the software project, and to determine the time periods when work elements of the WBS will occur.
Estimate the schedule as follows:
1. Allocate time for each work element of the WBS, and determine the work loading Allow at least one-month per year of fully-funded schedule margin; this is separate from any
Cost reserves. A recommended practice is to allocate the schedule margins at the timing of major reserves and/or transitions between life-cycle phases. For example, add one month schedule reserve per year after the PDR.
2. Determine the order in which work elements will be done. Define which work elements can be done in parallel, as well as dependencies that drive the schedule.
3. Based on the overall project schedule imposed on the software development, attack the scheduling problem from both ends. Start with the beginning date and create an activity network that shows the interrelationships between work elements. Then, start with the end date and work backward using the same activity network to see if the work elements integrate.
4. Determine the critical path through the schedule (longest path through the activity network in terms of time).
5. Smooth out the initial work loading to level non-critical path activities.
6. Inconsistencies and holes in the estimates may appear while scheduling the individual work elements and determining resource loading. This is especially true when trying to fit the work elements into the schedule imposed on the software project. As a result, it may be necessary to reiterate the estimates of other steps several times, to reduce the effort, or assume more risk to fit into the imposed schedule.
The outputs of this step are as follows:

- Assumptions made to estimate schedule
- Schedule including all work elements of the WBS, milestones, and reviews
- Revised estimates and assumptions made to revise estimate

Step 6 - Calculate the Cost

The purpose of this step is to estimate the total cost of the software project to cover the work elements and procurements of the WBS.
Estimate the total cost as follows:

1. Determine the cost of procurements:
Determine the cost of support and services, such as workstations, test-bed boards and simulators, ground support equipment, and network and phone charges.
Determine the cost of software procurements such as operating systems, compilers, licenses, and development tools.
Determine the cost of travel and trips related to customer reviews and interfaces, vendor visits, plus attendance at project-related conferences.
2. Determine the cost of training planned for the software project.
3. Determine the salary and skill level of the labor force.
4. Input the effort, salary levels, and cost of procurements into an institutionally supported budgeting tool to determine overall cost. All estimates should be integrated with all rates and factors, institutional standard inflation rates, and median salaries.
5. As with scheduling, inconsistencies and holes in the estimates may appear while calculating the cost. This is especially true when trying to fit the cost into the budget imposed on the software project. As a result, it may be necessary to reiterate the estimates of other steps several times, reduce the effort and procurements, or assume more risk to fit into the imposed budget. If the schedule becomes extended, costs will rise because effort moves out to more expensive years. See later steps for reviewing estimates versus budgets and schedule.

The outputs of this step are as follows:
• Assumptions made to estimate cost
• Methods used to estimate cost
• Cost of procurements
• Itemized cost estimates by WBS elements (in dollars)
• Total cost estimate (in dollars)

Step 7 - Determine the Impact of Risks

The purpose of this step is to identify the software project risks, to assess their impact on the cost estimate, and to revise the estimates based on the impacts.
The outputs of this step are as follows:
- Detailed software project risk list
- Assumptions made to revise estimates
- Methods used to revise estimates
- Revised size, effort, schedule, and cost estimates for risk

Step 8 - Validate and Reconcile the Estimate via Models and Analogy

The purpose of this step is to validate the estimates.
1. In addition to the main estimate that was developed in the preceding steps, obtain a second estimate, using one of the following:
a. Alternate Estimate
Have a second person or team, with similar software experience, generate independent estimates.
b. Historical Analogies
Using historical data compare the estimates with previous experience such as in the following areas:
- Size, effort, and cost of similar software
- Size versus functions
- Size versus effort and cost (development productivity)
- Technology versus effort and cost

2. Have the responsible people for this step meet to compare the main estimates with the second estimates, resolve the differences, and refine the estimates until a consensus estimate is reached. The lowest estimates should be given special scrutiny, as experience has demonstrated that estimates are usually low.
The outputs of this step are as follows:
• Assumptions made to validate the estimates
• Methods used to validate the estimates
• Validated and revised size, effort, schedule, and cost estimates with improved accuracy

Step 9 - Reconcile Estimates, Budget, and Schedule

The purpose of this step is to review the validated estimates with respect to the project-imposed budget and schedule and to resolve the differences. In many ways, Steps 9 and 10 are the most difficult steps in the cost estimation process, because of the need to understand, in an integrated manner, the cost of individual functions, their relative prioritization, and the functional interrelationships. If an inconsistency arises, there is a tendency to incorrectly address the issue as only a problem of incorrect estimation. However, in most cases, the real solution is to descope or reduce functionality, and then to descoed again, until the task fits the budget. Do not reduce costs by eliminating reserves and making optimistic and unrealistic assumptions.
1. Calculate the budget margin. Subtract the estimated cost from the budgeted cost. Then divide by the budgeted cost to get the margins. Multiply by 100 to get percent margin. Calculate schedule margin in the same manner.
2. Compare the estimated cost, schedule, and margins to the project-imposed budget, schedule, and margins to determine if they are consistent.
3. If the estimates are substantially greater, then identify and resolve the differences:

Refine the desired scope and functionality to the lowest level possible by analyzing and prioritizing the functions to identify

those functions that can be eliminated. Make certain you account for interrelationships between functions.
Begin eliminating procurements that are not absolutely necessary.
Revise the schedule, cost estimates, and risks to reflect the reductions in cost based on steps a-d. Reducing high-risk functionality or procurements can reduce risk and costs greatly.
d. Repeat the process until the functionality and procurements are affordable, with respect to the budget, and feasible, with respect to the imposed schedule.
Review the reduced functionality, reduced procurements, and the corresponding revised estimates with the sponsor to reach agreement. If agreement cannot be reached, higher-level management may need to intervene and assume a greater risk to maintain functionality. Update the WBS according to the revised functionality.
As the project progresses, it may be possible to include some functions or procurements that were originally not thought to be affordable or feasible.

The outputs of this step are as follows:

- Assumptions made to revise estimates
- Methods used to revise estimates
- Revised size, effort, schedule, and cost estimates
- Revised functionality and procurements
- Updated WBS
- Revised risk assessment

Step 10 - Review and Approve the Estimates

The purpose of this step is to review the software estimates and to obtain project and line management approval.
1. Conduct a peer review with the following objectives:

- Confirm the WBS and the software architecture.
- Verify the methods used for deriving the size, effort, schedule, and cost. Signed work agreements may be necessary.
- Ensure the assumptions and input data used to develop the estimates are correct.
- Ensure that the estimates are reasonable and accurate, given the input data.
- Formally confirm and record the approved software estimates and underlying assumptions for the project.

2. The software manager, software estimators, line management, and project management approve the software estimates after the review is complete and problems have been resolved. Remember that costs cannot be reduced without reducing functionality.
The outputs of this step are as follows:
• Problems found with the estimates
• Reviewed, revised, and approved size, effort, schedule, cost estimates, and assumptions

Step 11 - Track, Report, and Maintain the Estimates

The purpose of this step is to check the accuracy of the software estimates over time, and provide the estimates to save for use in future software project estimates.

1. Track the estimates to identify when, how much, and why the project may be overrunning or under-running the estimates. Compare current estimates, and ultimately actual data, with past estimates and budgets to determine the variation of the estimates over time. This allows estimators to see how well they are estimating and how the software project is changing over time.
2. Document changes between the current and past estimates and budgets.
3. In order to improve estimation and planning, archive software estimation and actual data each time an estimate is updated and approved, usually at each major milestone. It is recommended that the following data be archived:

- Project contextual and supporting information
  - Project name
  - Software organization
  - Platform
  - Language
  - Estimation method(s) and assumptions
  - Date(s) of approved estimate(s)
- Estimated and actual size, effort, cost, and cost of procurements by WBS work
- element
- Planned and actual schedule dates of major milestones and reviews
- Identified risks and their estimated and actual impacts
- The outputs of this step are as follows:
- Updated tracking comparisons of actual and estimated data
- Evaluation of the comparisons
- Updated size, effort, schedule, cost estimates, and risk assessment
- Archived software data, including estimates and actual

V **A**dvantage of **N**ew Phase in SDLC **M**odel

None of the SDLC models discuss the key issues like Effort Estimation, Cost Estimation management and Schedule management processes within the SDLC process, but, it is addressed in the overall project management.
The drawback of addressing these management processes under the overall project management is missing of key technical issues pertaining to software development process that is, these issues are talked in the project management at the surface level but not at the ground level.
This drawback in the existing SDLC model is rectified by the Software selection Phase in SDLC model by embedding the core control which is mapped to the project management and traceable to the surface level data existing to the project management and the ground level technical data additionally present to address Selection attributes such as Cost, Effort and performance and Schedule etc, issues during deployment of the software.

(A) In addition to this it has modified the concept of user-developer interaction in the conventional model to a three dimensional model which comprises of the user, owner and developer. It also establishes a clear guideline to address who has to do what in SDLC? And when in the different stages of

SDLC? So that the conflicts among the developer and user is removed also enables the timely completion of projects by removing the hidden overheads that occur due to conflicts or lack of knowledge.

(B) Very clear definition of roles and responsibilities ensures that conflicts among the developer and the user are eliminated or reduced to a manageable level. Hence, timely completion of software projects within the set timelines is made possible by removing the hidden overheads that arise due to lack of knowledge.

f) To ensure that the Selection attributes of a software project are adhered to, an outer circle has been added to ensure that these Selection attributes are embedded in the various activities and actions done by everyone involved in the development activities from requirement gathering to maintenance in various phases of SDLC.

VI. CONCLUTION

A major project management activity involves planning and estimating, and on IS projects developed using traditional approaches; the success of Selection has proved controversial to say the least. Reports of inaccurate Selection have typically cited customer change requests and unclear requirements as the main causes. Software selection is not an easy task but selections should always be reflective of the Software project Management. Existing SDLC life cycle not include a selection procedure in different phase of SDLC life cycle. In conclusion, research on Software Selection has been conducted for decades with immense quantities of models and tools produced. This study has looked at the Selection process in the emerging field of IT project development, and examined the causes of inaccurate estimates and steps to improve the process.

REFERENCES

[1]Raymond Lewallen - CodeBetter.Com - Stuff you need to Code Better! Published 08-01-2008.

[2] Craig Larman, Victor R. Basili, "Iterative and Incremental Development: A Brief History," Computer, vol. 36, no. 6, pp. 47-56, June 2003, doi:10.1109/MC.2003.1204375.

[3] www.coders2020.com/what-is-sdlc-what-are-the-various-sdlc-models-explain-them.

[4] Curtis, Krasner, Iscoe, 1988.

[5] Dr. Winston W. Royce (1929 - 1995) at www.informatik.uni-bremen.de. Retrieved 27 Oct 2008.

Svapnil Vakharia(ISTE-LM'10). he became a life member ship of ISTE in 2010. His date of birth is 9st September 1984 in kapadvanj. he completed B.E.Infromation Technology Engineer from Saurastra .University(Gujarat) with 61.53% 2006 , M.IS from Bllarat University(sydney,Australia) with 68.13% in 2008.

He is working as**, Asst. Professor in Gandhinagar Institute of Technology college, Gandhinagar.**

Usability Testing: Effective Testing Method To Improve Quality Of Software Design In Technical Education.

Birendrasinh.K.Zala, Gandhinagar Institute of Technology, Gandhinagar
birendra.zala@git.org.in

Abstract: **This paper will try to answer a question that how our foundation of software designers can transform from good programmers to good User Interface Designers. To make this happen, this paper tries to enlighten on a very efficient but not very used testing method in our country called Usability testing for next generation of software developers produced by our technical education. This paper will discuss all steps of usability testing in a very general manner to ensure that fundamental thinking of them about software design and testing gets improved and very much clear without any complexity and confusion. The research method used in this method is evaluation of prototype.**

Index terms: Software Development Life Cycle (SDLC), User Centred System Design (UCSD), User Interface (UI), Human Computer Interaction (HCI)

Keywords: Prototype, user involvement, Usability Testing

I INTRODUCTION

Everyone just wants to become a programmer. No one cares about how good design of user interface they have produced in terms of interaction style, complexity, functionality. Especially in case of technical engineering colleges, whose responsibility is to produce good software engineers instead they are only focusing on just producing programmers. Why? Simple reason is that they just don't know importance of design and not even meanings of these words. This clears the idea that our next generation and also previous generations of software developers are not software engineers rather than they are software programmers only. Operational product should be produced that is the only thing our programmers want because of a myth that what user wants is operational product and everything else is secondary things. In terms of who is responsible for this mentality in their minds is our technical education which doesn't provide any motivation for UI designers by designing of content of courses which mainly focus on the subject of programming. In terms of effective design of software, they only rely on the traditional methods like Unit testing, Integration testing etc…Even they are not using these testing methods to test their applications because who cares how effective UI they have produced. Even though they use these methods still they are not very best to identify the effectiveness of UI. And what is the end result of all these scenarios, failure of design of interface. Here Human Computer Interaction (HCI) comes in light.

HCI focus on verifying three main things like system is self learnable for users [1], complexity of interaction and filling the gap between the system image and user' model. To get success in all three factors we don't need to look around we already have one approach called User Centred System Design (UCSD) [2]. There are certain guidelines of UCD we need to follow during the design of UI to ensure a proper match between them. User Centered Design is a design approach, adopted by many industries in order to develop products and services that will meet the needs and expectations of the end users [3] .One of the guidelines includes testing of UI by testing method called Usability Testing.

This paper discusses a very well developed testing method called Usability Testing to check whether UCD approach is followed in Software design or not. This method includes System Description activity, System Redesign activity and Evaluation activity.

- System Description phase gather, identifies and specifies all basic requirements to produce a base for system redesign phase.
- System Redesign phase focus on redesigning the UI on based on the facts gathered in System Description phase.
- System Evaluation phase includes activities of testing the prototype which was produced during System Redesigning phase.

II USABILITY TESTING PHASES

All three basic phases of Usability Testing will be explained in this part of paper with respect to one example system which is a Blackboard tool for a university they are developing for the faculty members and students.

➔ *SYSTEM DESCRIPTION PHASE*

In this phase as a system analyst you have to convince your client that you understand the system domain, the users and their context of use. To get approval from client you must show evidence that you:

• Define all problematic areas of the system on which a user's navigational capabilities depend on and which must be improved to increase their satisfaction for using application.
•You have collected data from user interviews and/or observations.

Deliverable of System Description phase is System Description report which includes:

- ✓ System Description.
- ✓ User Description.
- ✓ User goals.
- ✓ Task Description.
- ✓ The Users' Environment.
- ✓ Analysis of current system.
- ✓ Concrete use cases for task of interest.
- ✓ Usability metrics to define success during Evaluation.

The table 1 shows the concrete use case format where user action and system response will be identified to draw a use case.

Table .1 Concrete Use Case

	USER ACTIONS	SYSTEM RESPONSE
1		
2		

➔ *SYSTEM REDESIGN PHASE*

➔ System redesign ensures that a suitable design should be made which focus on improving the defects uncovered during description phase. This redesign is intended to be a reflective exercise based upon the theory and usability principles and guidelines.

Outcome of System Redesigning phase is System Redesign report which includes:

- ➢ Creation of use scenario that describes how the redesigned user interface will support user's selected task which explains how the new design will be an improvement upon the original.
- ➢ Creation of an essential use case for the use scenario for justification of how the different tasks have been allocated between the user and the system.
- ➢ Create a concrete use case corresponding to the essential use case. Mark up the concrete use case, indicating the task objects and attributes.
- ➢ Create a detailed content diagram for the part of the system you described in your essential use case. Explain the content diagram.
- ➢ Prototype will be produced as a main deliverable and will serve as a foundation for Evaluation phase. Figure 1 shows a one web page paper prototype [4] design for Blackboard tool.

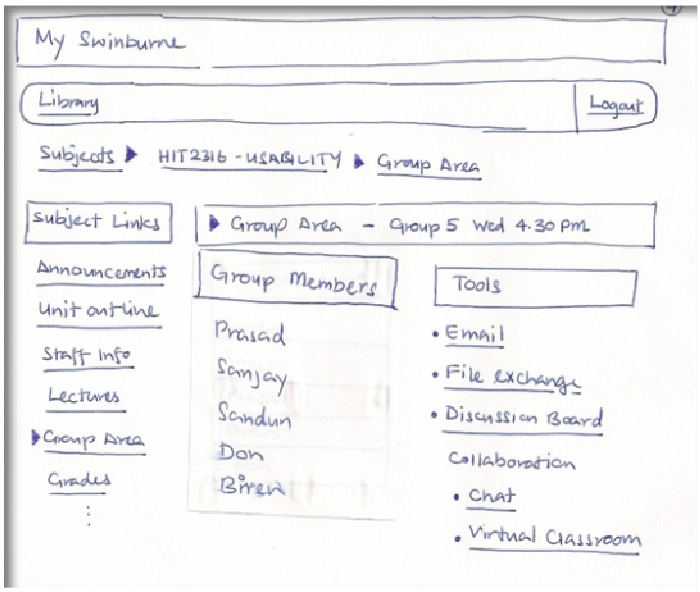


Figure 1: Paper Prototype Design for Blackboard tool

➔ *SYSTEM EVALUATION PHASE*

System Evaluation this phase mainly differs from other testing methods .It includes activities like:

➔ Create a task description for use in your usability test. Justify the task/s to be carried out.

- ➔ List the role each team member will play in the evaluation. There will be main four roles we will require our team members to perform like Observer, Product Expert, Facilitator and Test Administrator.
- ➔ Prepare demographic questionnaire and any extra questions you plan to use in the evaluation..
- ➔ Prepare a copy of observation sheet.
- ➔ Gather and analysis of data gathered during testing.
- ➔ Calculate metrics to identify efficiency and effectiveness of UI which is shown in table 2.

Table 2 shows the metrics for calculating the results of particular task by each participant to calculate effectiveness and efficiency.

Sr no.	Unassisted Task Completion Rate	Errors	Task Completion Time(mins)
P1			
P2			
P3			
P4			

Table 2: Task Metrics for Individual task

III BENEFITS AND DRAWBACKS

Here are some of the benefits we can get from usability testing:

- ➔ Preventing Usability Problems in the First Place [5].
- ➔ As user and user needs are at centre point of focus throughout the testing, probability of resultant design will be user oriented tends to increase more than ever which is most important for a application to be a successful.
- ➔ The user can experience the look and feel of the [6] future system which leads to proper implementation of user's model in their mind.
- ➔ It increases the active user involvement which reduces the risk of changing requirements.

Drawbacks of usability testing are as follows:

- ➔ It is more time and cost consuming than other testing methods.
- ➔ It demands complete understanding of users and their task they need to perform to design task description and test cases otherwise results can mislead the development.
- ➔ Special equipments require which adds more cost to original budget.
- ➔ It demands knowledge of all UCSD standards and principles from designer side to effectively conduct test.

IV CONCLUSION

Usability test conducts on paper prototypes to identify a proper match between systems image and user's model to increase the chance of success of application .Additionally this paper tries to prove a myth wrong that that usability testing is a big production [7] because usability testing doesn't need to be perfect one. As it focuses mainly on gathering data from user don't worry about finding users, you can just starts with your co-worker.

V REFERENCES

[1] C. Abras, D. Maloney-Krichmar, and J. Preece, "User-Centered Design," In W. Bainbridge (Ed.), Encyclopedia of Human-Computer Interaction, Thousand Oaks, Sage Publications, 2004.

[2] J. Gulliksen, B. Goransson, I. Boivie, S. Blomkvist, J. Persson, and A. Cajander, "Key principles for user-centred systems design," In Behaviour and In- formation Technology, 22(6), 2003, pp. 397-409.

[3] J. Heinila, H. Stromberg, J. Leikas, V. Ikonen, N. Iivari, T. Jokela, K. Aikio, I. Jounila, J. Hoonhout, and N. Leurs," User Centred Design Guidelines for Methods and Tools," The Nomadic Media consortium, 2005, pp. 5.

[4] S. Sade, M. Nieminen, and S. Riihiaho," Testing usability with 3D paper prototypes—Case Halton system," Applied Ergonomics, 29 (1), pp. 67-73, 1998.

[5] J. M. Spool (2007, Feb 27)" Usability Testing: Oh, The Things You Can Learn," UIEtips. Retrieved March 16, 2011. from http:// www.uie.com /articles/usability_tests_learn/

[6] A. Lancaster," Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces," IEEE Transactions on Professional Communication, 4, 2004, pp. 335-336.

[7] C. Perfetti, (2008, Sep 02), "Five Techniques for Getting Buy-In for Usability Testing," *UIEtips*, Retrieved March 16, 2011, from http://www.uie.com /articles/ usability_buy_in_reprint/

VPN based on Traditional and MPLS

Rahul A Vaghela, Asst. Professor,
Gandhinagar Institute of Technology,
rahul.vaghela@git.org.in

Abstract: **The Virtual Private Network has become almost as wildly used in the networking industry as has Quality of Service to describe a broad set of problems and solutions. The idea is to create a private network via tunneling through the encryption over the Public Internet. Of course it's a lot cheaper than using own frame relay based connection but again it seems to be sticking cotton in your ears in a Hall and pretending nobody else is around.[6] This paper will discuss the two different approaches for implementing VPN. The Traditional approach will be a best effort networking in which traffic treated equally and packets can be discarded upon congestion. While the MPLS approach combines the concept of routing and switching.**

I. INTRODUCTION

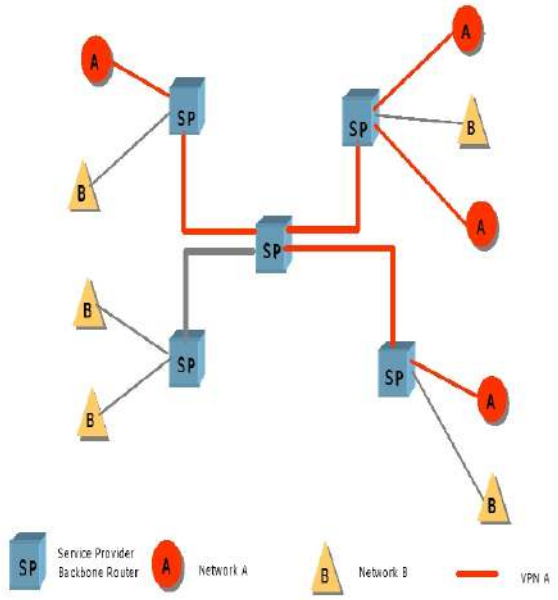
The simplest definition of VPN can be described by the meaning of the each word in the acronym individually and then subsequently tie each of them together in a common and meaningful fashion. The meaning of the word network is very wide and can be describe as set of devices having communication of some arbitrary method. Devices may include computers, printers, routers and so forth and they may reside in geographically diverse locations. They can communicate with numerous methods. For simple, it is a bunch of devices that can successfully transmit and receive data amongst themselves. The Term Private fairly clear and can be said as if two people are communicating then the content of their communication kept secret or confidential.

'Virtual' is a concept which is slightly complicated to describe. The dictionary meaning of it is the imaginative substance and the antonym is the real. In short you are going to operate on the substance which is imaginary.

Altogether the private communication conducted across the network infrastructure that is shared by multiple organizations and the virtualization of private network is a network with no physical counterpart and the participants not taking part in communication are not privy to the content of data as well as the virtual relationship of the terminals involved in the virtual communications. It can be built between two end systems or two organizations or multiple users in single organization or multiple organizations across the global Internet. The formal definition of VPN is:

"A VPN is a communications environment in which access is controlled to permit peer connections only within a defined community of interest, and is constructed though some form of partitioning of a common underlying communications medium, where this underlying communications medium provides services to the network on a non-exclusive basis."[6] In simpler, the VPN is private network constructed within a public network infrastructure such as global Internet

As shown in figure, Network 'A' sites have established a VPN across the backbone of he network which is depicted by red lines, where the Network 'B' is completely unaware of its existence. Both Network A and Network B harmoniously coexist on the same backbone infrastructure.

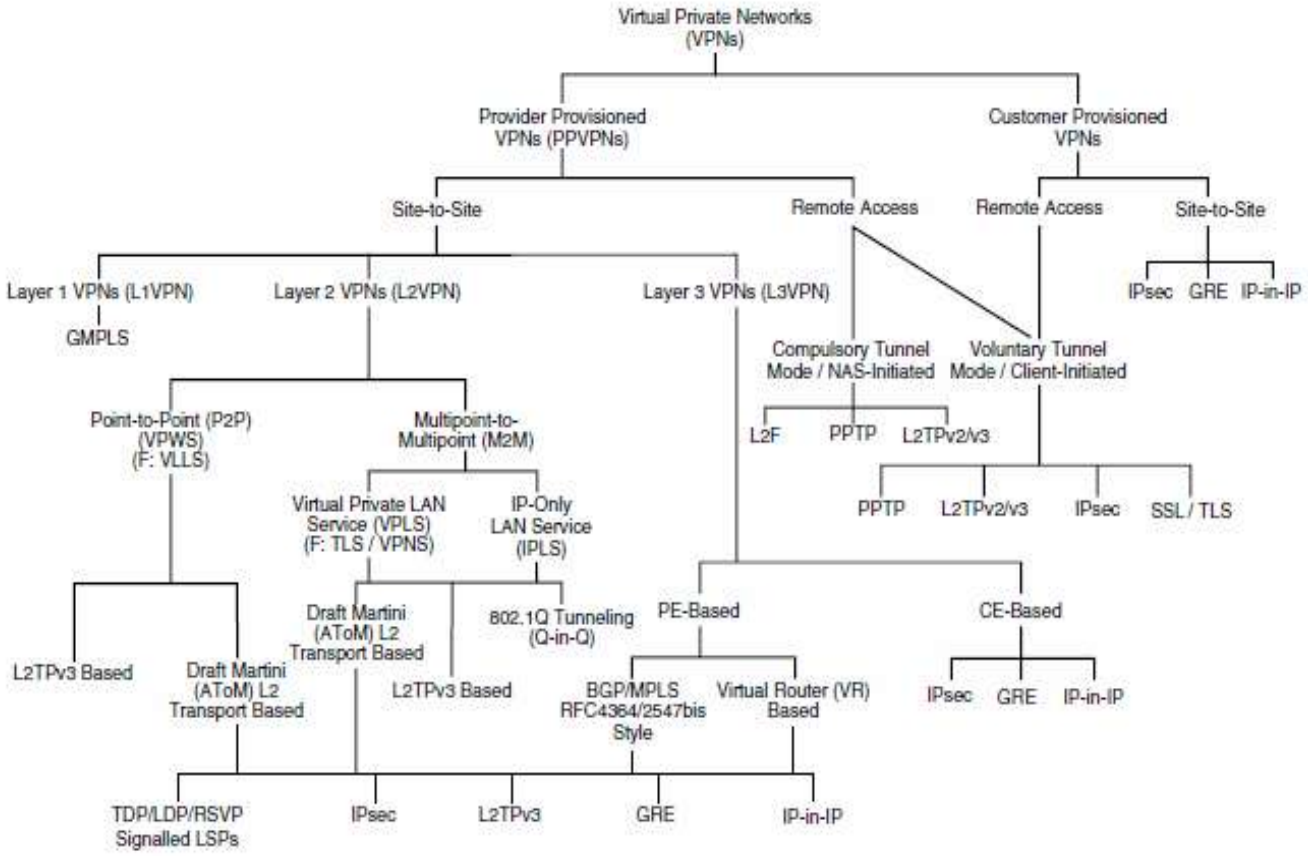


II. VPN DEPLOYMENTS

There are actually different types of VPN available based on the functional requirements and the methods applied to build them. The selection process depends on what consideration has been taken to solve the problems, risk analysis of the security provided by the specific implementation, how much expansion can be possible in terms of size and the complexity involved in implementation, maintenance and troubleshooting

A. Traditional Routing

The traditional IP network consists of a series of routers interconnected by physical media that communicate via standard routing protocols. The need for robust communication was one of the goals of IP networking in its early days. Delivery of packets with specific delays or



bandwidth requirements was not an issue. Even though IP has the concept of type of service it has rarely been utilized. IP has been a very powerful networking technology. Conforming to open standards and its flexibility enables it to transfer a wide range of data types. [4]

The growth of the Internet has put IP to the forefront of the communication world. The Internet is generally segmented into many autonomous system domains and uses typically an interior gateway protocol such as OSPF to route packets inside the AS and an exterior gateway protocol such as BGP is used to communicate between routes of separate AS. Traditional routing being Connectionless has some benefits in terms of scalability and network resiliency. In a service provider network this connectionless nature also has some limitations. [6]

Within the network OSPF establishes links using the open shortest path first algorithm. Some of the difficulties encountered are a likelihood of congestion on some of the links and a limited ability to distribute traffic over all available links. One other issue to consider is that traffic is sent across links on a hop-by- hop basis. Routing decisions are made at every node. This can create congestion on the network because routers base their forwarding decision on the destination address on the packet header and the least path cost to that destination. This causes all traffic to that destination to take the least cost path leaving the other links underutilized. With traditional routing service providers can provide only a best effort networking. All traffic is essentially treated equal and packets may be discarded on congestion. This is acceptable for application like e-mail and others with no specific requirements for latency or bandwidth.

A. Need for Switching

As service provider networks grew large the problems of the forwarding component of traditional routing and the difficulty of predicting performance in a large meshed network grew. Many service providers have enhanced their IP services by incorporating ATM and Frame-Relay, a connection oriented service into their networks. ATM and Frame-Relay networks use a different forwarding method based on a label-swapping algorithm. This forwarding is done in hardware and yields a much better performance than traditional routing. Frame-Relay and ATM are connection-oriented technologies.

A connection is established between the two end points before the traffic is sent. Since the traffic is sent along a predetermined path the network becomes more predictable. When data is sent through an ATM network the end-to-end connection takes into account the state of the network, the latency and bandwidth requirements of the application and the preferred routes. These are critical to ATM's connection oriented performance.[1]

Many large networks now have a switching fabric at the core. IP routing still dominates at the network edge. The need to overlay and integrate the best of traditional layer 2 and layer 3 technologies raise some interesting issues. When an IP network is overlaid on top of an ATM switched network all the

routers seem to be connected at the network layer. This requires every router to have an adjacency with every other router. The adjacency must be established using ATM Virtual Circuits (VC).[1] As the network expands, the VC's requirements grew exponentially limiting the scalability of the network. This is generally referred to as the "n-squared" problem. One other issue on the ATM network is the 5-byte overhead on every 53-byte ATM packet and the difficulty in performing very high speed segmentation and reassemble (SAR). Current IP networks are still far from meeting all the requirements of service providers and their customer

B. Multi Protocol Label Switching

MPLS is designed to meet the mandatory characteristics of a large-scale carrier class network. It uses existing layer 3 routing protocols as well as the widely available layer 2 transport mechanisms and protocols. The IETF set up the MPLS working group in 1997 to develop a common standardized approach. The goal of the MPLS working group was to standardize protocols that used Label Swapping forwarding techniques. The use of label swapping has powerful advantages. It separates the routing problem from the forwarding problem. Routing is a global networking problem and requires the cooperation of all anticipating routers. Forwarding is a local about to take problem. The router/switch decides entirely on its own which path MPLS has one more advantage. It reintroduces the Connection State into the IP data flows.

MPLS integrates the best of layer 2 and layer 3 technologies. The key component within a MPLS network is the label switching router (LSR), which is capable of understanding and participating in IP routing, and layer 2 switching. MPLS has provided significant new capabilities in four areas that have ensured its popularity: QoS support, Traffic Engineering, Virtual Private Network, and Multiprotocol Support.

QoS is the ability to assure delivery of important data flows. Network Mangers require QoS for many reasons. Some of the requirements are

1) Guarantee a fixed amount of bandwidth for various applications
2) Control latency.
3) Provide quantifiable SLA.
4) Ability to configure various levels of QoS for multiple customers. [1]

In a connection less IP based internetwork, it is very difficult to provide any true Qos commitments. A Differentiated Service (DS) or Integrated Services (IS) with RSVP is limited in terms of flexibility and scalability and may prove inadequate in a heavily loaded network. A connection-oriented service has powerful traffic management and QoS capabilities. MPLS imposes a connection-oriented framework and provides the foundation for reliable QoS traffic contracts.

Traffic Engineering is the ability to dynamically plan resource commitments on the basis of known demands, define routes dynamically and optimize network utilization. Traffic engineering seeks to control traffic flows and network resources so that predefined objectives can be met. With basic IP routing there is a primitive form of automated traffic engineering. Dynamic routing reacts in a very simple manner to congestion and does not provide a way to support QoS. When MPLS is applied, the layer 2 circuits are replaced by Label Switched Path (LSP). A set of protocols and tools are designed to measure traffic [1]

MPLS provides an effective mechanism for supporting VPN's. MPLS technology provides the ability to separate traffic belonging to different VPN's. In addition the establishment of LSP tunnel satisfies the formation of a virtual topology. One more advantage of MPLS as a VPN tunnel technology is MPLS traffic engineering can dedicate resources to a LSP. The security of a VPN tunnel using MPLS is equivalent to that provided by ATM/Frame-Relay PVC.

MPLS can be used on many networking technologies. MPLS enabled routers can coexist with ordinary routers. MPLS is designed to work in ATM, Frame-Relay networks. MPLS enabled ATM, Frame-Relay switches can also work with ordinary switches.

C. MPLS Operation

The MPLS network consists of a set of nodes capable of switching and routing on the basis of label appended to each packet. A MPLS domain consists of a contiguous or connected set of MPLS enabled nodes. These nodes are called Label Switched Router (LSR). The labels define the flow of packets between the two endpoints. A specific path through the network of LSRs for each distinct flow called a Forwarding Equivalence Class (FEC) is defined. MPLS is a connection-oriented technology. With each FEC is associated a traffic characterization that defines a QoS requirements for that flow. Because the LSR forwards the packet based on its label value this ensures that the forwarding process is simpler than with an IP router.

Figure 1: MPLS Operation

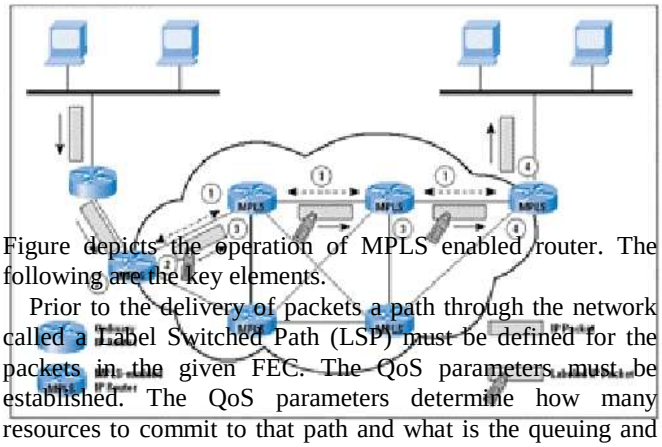


Figure depicts the operation of MPLS enabled router. The following are the key elements.

Prior to the delivery of packets a path through the network called a Label Switched Path (LSP) must be defined for the packets in the given FEC. The QoS parameters must be established. The QoS parameters determine how many resources to commit to that path and what is the queuing and discard policy at each LSR for the FEC. To accomplish the

above an interior gateway protocol like OSPF is required for reach ability and routing information. Labels are assigned to each packet in the FEC. These labels have only local significance. A protocol like Label Distribution Protocol (LDP) or an enhanced version of RSVP is used to determine routes and establish labels values. These can also be setup manually by an operator.

A packet enters the MPLS domain through an ingress edge LSR. It is here that the packet is processed to determine its need for network layer services, defining its QoS. The LSR assigns this to a particular FEC and therefore a LSP, and then forwards the packet. Each LSR that receives a labeled packet removes the incoming label and attaches the appropriate outgoing label to the packet and forwards the packet to the next LSR in the LSP. The egress edge LSR strips the label reads the IP packet header and forwards the packet on to its final destination. [1]

One of the most important features of MPLS is label stacking. A labeled packet may carry many labels organized as a last-in-first-out. Processing is based on the top label. At any LSR labels can be added to the stack or removed from the stack. This allows aggregation of LSPs into a single LSP for the portion of the route creating a tunnel.

The FEC for a packet can be determined by one or more parameters, such as source or destination IP address, source or destination ports, IP protocol ID, differentiated service codepoints or IPv6 flow labels. A per-hop behavior (PHB) can be defined at a LSR for a FEC. The PHB defines the queuing priority of the packets for this FEC and the discard policy. Packets send to the same endpoints may belong to different FEC and will be labeled differently and experience different PHB at each LSR and may follow different paths through the network. The essence of MPLS functionality is that traffic is grouped into FECs.[1] The traffic in a FEC transits a MPLS domain along a LSP. Individual packets in a FEC are uniquely identified as being a part of a given FEC by means of a locally significant label. Route selection refers to the selection of a LSP for a particular FEC. MPLS supports hop-by-hop routing and explicit routing. With hop-by-hop routing each LSR independently chooses the next hop for each FEC. This option makes use of an ordinary routing protocol such as OSPF. This has some advantages but because of the limited use of performance metrics, hop-by-hop routing does not readily support traffic engineering or policy related to QoS and security. With explicit routing a single LSR specifies some or all the LSRs in the LSP for a FEC. [1]

Explicit routing provides all the benefits of MPLS, including the ability to do traffic engineering and policy routing. Dynamic explicit routing provides the best scope for traffic engineering. In this mode the LSR setting up the LSP would need information about the topology as well as QoS related information for the MPLS domain. The enhanced version of OSPF for MPLS has some newer metrics that would be useful in constraint based routing including maximum link data rates, current capacity reservation, packet loss rate and link propagation delay. Route selection consists of defining a LSP for a FEC. A separate function is the actual setting up of the LSP and for this each LSR on the LSP must

a) Assign a label to the LSP to be used to recognize incoming packets that belong to the corresponding FEC.
b) Inform potential upstream nodes of the label assigned by this LSR to this FEC.
c) Learn the next hop for this LSP and the label that the down stream node has assigned to this FEC.[1]

III. CONCLUSIONS

In conclusion, The MPLS provides all the required benefits needed by Service provider to provide networks like predictability, scalability and manageability. It defines the revolutionary methodology in combination with layer 2and 3 technologies and in core having the concepts of switching and routing technology at edge. It provides better routing and forwarding scalability and traffic engineering capabilities for better network provisioning. It would be the excellent choice for deploying the VPN services.

REFERENCES

[1] Sinha R, *"MPLS –VPN services and Security"*.,GSEC Practical version 1.4b, May 2003
[2] Huang W., Yang, J, New VPN Application in 3G Network, ISBN 978-952-5726-07-7,2009
[3] Wagner D., Analysis of SSL 3.0 Protocol, Http://www.Counterpane.com,2005-08-07.
[4] Rosen E, Viswanathan,R ,"A Proposed Architecture for MPLS," draft-ietf-mpls-arch-01.txt, March 1998.
[5] Kolar T, Valencia A, Littlewood M, *"Layer Two Forwarding (Protocol) 'L2F'*," draft-valencia-l2f-00.txt, , October 1997.
[6] Heinanen, J, Rosen E., "*VPN Support for MPLS*," draft-heinanen-mpls-vpn-01.txt,, March 1998.
[7] Frguson P, Huston G, *"What is a VPN"*, April,1998

Rahul A Vaghela is a life member of Indian society of technical Education and his membership number is LM 69699. The author has completed his Bachelor of Engineering in Information Technology from Gujarat University. Then after, He completed the Master of Information Technology from Australia in the year of 2007. He worked as Software Engineer for Jav IT Pvt Ltd for three years in Australia. He joined the Gandhinagar Institute of Technology in the year of 2009 which is approved by AICTE and affiliated to Gujarat Technological University. Currently the author is working as the Assistant professor and also Head of Information technology department.

Application of composite material of Carbon fibre and Glass fibre with Nylon 6 on industrial product

[1]Prof. A.D. Patel, [2]Asso. Prof. B S Patel, [3]Asst. Prof R S barot

Mechanical Engineering Department, BVM Engineering college,
Vallabh Vidyanagar, Anand, Gujarat, India.

[1]astorcan@hotmail.com,[2]bharatvimlapatel@yahoo.com,

[3]rakesh_le8@yahoo.com

Abstract

In a pneumatic accumulator the requirement and functional aspect for the piston is a high strength, high stiffness, low weight, tough and able to withstand impact loading, durable and does not creep or warp as a result of exposure to temperature or humidity changes, and can be processed into the required shape. The requirement for high strength and low weight can be translated into a requirement for a high value of strength/density, i.e. specific strength. Possibilities would seem to be metals, polymers and composites. C I piston has the advantages that it is tough, has good specific strength, well damping properties for vibrations and is cheap. The specific stiffness could be better. Aluminium alloys have the advantages of toughness and good specific stiffness. They are, however, more expensive than cast iron. Another problem is that they have very poor vibration damping. Aluminium can be protected against corrosion attack by damp environments by anodizing. An aluminium piston could be made by casting into the required shape and machining.

We had developed composite material of nylon 6 polymers with carbon fibre and glass fibre to increase mechanical properties for designing piston of pneumatic accumulator. We have used Twin Screw extruder (42 mm diameter, parallel counter-rotating twin screw with l/d=6.5 single feed port) and injection moulding process for manufacturing of tensile test sample according to ASTM D-638.The composite made by adding carbon fibre and glass fibre with nylon 6 polymer increase the tensile strength . We have found that the strength values are similar for samples containing the milled and chopped carbon fibres.

Introduction:

Day by day the demand for new materials is being generated to produce more complex products which result from technological break - through and the wish to obtain improved cheaper materials for existing application. The concept for developing new materials or composites, which should give optimum result and accelerated engineering properties, for this purpose, with the development of new methodology we have to concentrate on materials, but now the trends is coming for composites materials to achieve engineering design and manufacturing concepts.

We apply the concept of Value Engineering for Design and Development of Piston for pneumatic accumulator

Applying the value engineering concept on pneumatic accumulator at Cipriani Harrison valves private Ltd. Located at Vitthal Udyognagar. Dist. Anand. Gujarat.

We apply the concept of improved design and material change will result in high performance and cost reduction.[1] Value Engineering concept is apply to improve quality of product and cost reduction by applying material and design change of PISTON used in pneumatic accumulator. [2] Processing of FRP/Composites involves complex chemical reactions Final properties are determined by many factors including the type, amount, and composition of there in systems and reinforcements .In addition, the use of additives can greatly affect the FRP/Composite properties. Benefits of There can be many benefits obtained by the use of FRP/ Composites. These benefits and characteristics should be considered early in the design process.

STEPS IN PRODUCT DESIGN: By definition, the optimum design of a mechanical element is the selection of the material and the values for the independent geometrical parameters with the objective of either minimizing an undesirable effect or of maximizing a functional requirement, making certain in the design procedure that the element satisfies other functional requirements and that other undesirable effects are kept within the tolerable limits. From the data collection, The Company are using piston material of Aluminium alloy of different combination, 2.24 % Mn-Al alloy and 1.25 % Mn- Al alloy.

By applying Johnson method of optimum design of a mechanical element is the selection of the material. [3]

JOHNSON'S METHOD OF OPTIMUM DESIGN

In Johnson's method of optimum design, any mechanical element design can have three forms of equations;

a). Primary Design Equation.(PDE)

This is the most important design equation, which expresses the most significant functional requirement or the most significant undesirable effect.

PDE (Cost of material is the criterion of Optimization)

Let,

c_m = Cost of the bar in Rs.

p = Mass density in Kg/m^3

A = Area in m^2

L = Length in m

c = Cost of material / weight in Rs/N

c_m = Weight of the bar (N) x c = (Mass Density x Volume) x g x c

= p A L g c

PDE is, c_m = c p A L g

b). Subsidiary Design Equations (SDE)

In a mechanical element, design equations other than PDE are called Subsidiary Design Equations (SDE). SDE express either functional requirements or significant undesirable effects. The most important SDE are the stress equations, which are generally implied. SIDE

$$\sigma = \frac{P}{A} \quad \ldots. (b)$$

c). Limit Equations or Constraints (LE)

Various parameters have certain ranges which are expressed by simple mathematical limit equations. For example, the limitations in stress are imposed by the strength of the material. Limitations on geometry are imposed by certain functional requirement.

The permissible stress should be less than or equal to

$$\frac{\text{Yield Strength}}{\text{Factor of Safety}}$$

Basic Steps in method of Optimum Design are:

1. Decide on basic configuration in the form of sketches, identifying loads, deflections and stresses.
2. Summarizing the significant constraints to compile the Primary Design Equation, Subsidiary Design Equation and the Limit Equation.
3. The equations are so combined that the total number of equations is equal to the number of free variables.
4. These equations are then studied for optimum design[4].

The above procedure would be clarified in the following sections. Three types of problems are encountered in the method of optimum design :

i). Case of Normal Specifications (NS)

ii). Case of Redundant Specifications (RS)

iii). Case of Incompatible Specifications (IS) the material selection factor (MSF) will be :

$$(MSF) = \frac{cp}{S_y}$$

The cost would be minimum for the material having lowest value of (MSF). From the data given, calculating (MSF), it is seen that 30% Glass filled nylon has the lowest MSF and hence should be used to achieve optimization in cost for given conditions.

Table : Material Selection Factor [MSF]

Material	Mass Density in Kg/M^3	Cost of Material in Rs/Kg.	Yield Strength [MPa]	MSF=c/Sy
2.24% Mn Aluminum alloy	2700	600.00	55	29454.54
1.25% Mn Aluminum alloy	2800	700.00	57	34385.96
30 % Glass filled nylon	1400	225.00	11.5	27391.30

We found from calculation for Material Selection Factor (MSF). For the first alloy i.e. 2.24 % Mn-Al alloy it is 29454.54 and for second alloy i.e. 1.25 % Mn-Al alloy it is 34385.96. And we suggested new material was 40% Glass-fibre reinforced polyamides has MSF was 27391.30. It shows that our suggested new material is most suitable for piston in pneumatic accumulator.

Applying the concept of Value Engineering and Value analysis, the Polymer piston is developed under the DARSIRI Method:

D = Data collection: Gathering information related to the piston used in accumulator

A =Analysis: developing new ideas for performing function of piston.

R =Records of ideas: Recording the minutes of analysis of Piston Design in meeting.

S =Speculation: Covering the ideas into practical suggestions for piston material and development of polymer composites piston.

I =Investigation: Testing the trial for performance of piston is to be carried out in the industry and check the working of piston.

R = Recommendation: Deciding which suggestion gives the best value for the working of piston.

I = Implementation: Assembling the piston of glass-fibre with all parts of Pneumatic accumulator.

PISTON DESIGN:

The piston is a disc which reciprocates with in a cylinder and is either moved by pneumatic air which enters the cylinder. The following points are to be considered in the design of pistons.[5]

1. Strength to resist the pneumatic air and inertia force.

2. Dispersion of heat, if any.

3. Pneumatic air sealing of the cylinder.

4. Bearing area sufficient to prevent undue wear.

5. Minimum weight.

6. Noiseless operation.

7. Resistance to mechanical distortion.

8. Adequate support for a piston pin.

The design of a piton and its dimensions are determined by the mechanical loads, the shape and arrangement of cylinder, the number of piston rings, the present of reinforcing ribs, the thickness of the piston crown.

HELICAL SPRING WITH AXIAL LOADING

Helical spring consists of a wire or rods wound about a mandrel to form a helix and is primarily intended for axial direct compression or tension loads. A helical spring is considered close and coiled if the plane containing each coils is nearly perpendicular to the axis of the helix. This will be the case when the helix angle is small and generally a spring is taken as close coiled if the angle be less than 10^{o}.

The design of any spring must satisfy two requirements:

(a) The spring must carry the service load without the stress exceeding the safe value.
(b) The force deflection characteristics, i.e. spring rate must be satisfactory for the given application.

After designing piston and helical spring, We make drawing of piston for new material, Glass fibre polymer with all design aspect with positive consideration.[6]. And according to this design, We developed the mould for polymer piston..After estimation of mould cost.. We develop mould for piston. Mould cost is consist of raw material cost of Rs.9088/-, Machining cost of Rs.7900/- Designing and developing cost of piston is Rs.3000/-, Heat treatment cost of mould is 1000/-, Assembly cost of mould is Rs. 500/- and Overhead cost for mould manufacturing is Rs. 500. So the total cost of mould is Rs.21998/- (Approximately Rs. 22000/-)

After manufacturing mould, it is send to Injection moulding company for piston to mould. We had check piston for all quality level and made assembly of Pneumatic

Accumulator and get all satisfied testing results.

Table

Material	Mass density In Kg/M^3	Volume of piston in mm^3	Weight of Piston in Kg.	Cost of Material in Rs/Kg.	Cost of Piston in Rs.
2.24% Mn-Aluminium alloy	2700	129	0.350	600.00	237.00
1.25% Mn-Aluminium Alloy	2800	85	0.230	700.00	186,00
30 %Glass filled nylon	1400	163	0.180	225.00	100.00

Conclusion:

We concluded from my product development project, the new polymer piston is saving 57.81 % of cost as compare to Al- alloy piston. Also the change in Design of piston will reduced the weight of piston from 0.350 Kg. To 0.180 Kg. i.e.51 % weight reduction, which leads to us for cost of spring to be reduced and this will give reduction of helical spring cost from Rs. 145/- to 120/- i.e.18% cost is lower than spring they were using before our product study.

From my studied on different engineering polymer, only one polymer cannot give all require characteristic properties, but by adding filler material like Glass fibre will increase the mechanical strength of polymer. The different composites developed by reinforcing fibre material with engineering polymer like Polypropylene, ABS and Polyamides, we conclude that for the development of glass fibre composites with Polyamides (Nylon-6) will increased the mechanical properties like tensile strength as we increased the amount of filler material. Also this product development leads to ease of manufacturing of piston, it can be made to have a long life, excellent resistance to corrosion and this polymer material is recyclable, degradable and performs all technical working conditions. Thus we try to reach versatile and multifunctional aspect of industrial products. Which may leads us towards growth of industries worldwide.

References

1. Diamond, W. J. Practical experimental designs. Munich:Hanser.1988.

2. Faupel, J. H. Engineering Design. A synthesis of stress analysis and materials Engineering. 2 ed. New York: Wiley, 1981.

3.. Rosato, D. V. " Material Selection." Encyclopaedia of composites. S. Lee.Ed. New York,VCH. 1990.

4. Boyer, H. E. Selection of Material for Service Envionments."Source Book, Metal Park, OH.ASM Int.1987.

5. Product Design:" Basic Processing guide for plastics. " SEP ANTEC,pp,1646-50,May, 1989.

6. Tong, L., Mouritz A. P. & Bannister M. Fiber reinforced polymer composites (Elsevier science, Oxford, 2002

Review of Boiling Heat Transfer Correlations for Refrigerants R 22 and R 134a

Nimesh Gajjar[1], N M Bhatt[2]
[1,2]Gandhinagar Institute of Technology, Gandhinagar
nimesh.gajjar@git.org.in

Abstract

Recent VCR cycle based systems mainly uses R134a as a substitute for R12 which was widely used in refrigerator, water cooler, car air conditioner etc. Problems of ozone depletion and global warming had compelled us to phase out CFC refrigerants. R22 has many domestic and commercial applications like window air conditioner, central air conditioning systems, deep freezers etc. and being HCFC, it has also replaced R11 and other CFC refrigerants for many applications like central air conditioning systems, deep freezers etc. This paper presents a review of different correlations available in literature for boiling heat transfer of refrigerants HFC134a and HCFC22 used in VCR cycle based equipments. From these correlations we can obtain evaporative heat transfer coefficient and average Nusselt number of refrigerant by implementing non-dimensional parameter so; better performed refrigerant can be selected from the perspective of initial cost of heat exchanger and running cost of compressor. The correlations are reviewed for different types of tube surfaces like smooth, rough, finned and grooved; and for horizontal and vertical orientation. So, one can choose suitable refrigerant which has higher heat transfer coefficient and higher COP for their equipment. Higher heat transfer coefficient will be reduced initial cost of the equipment and higher COP for the refrigerant will lead to lower power consumption.

Nomenclature

A - Cross sectional area (m^2)
Bo - Boiling number = $q/ G.h_{fg}$
Co - Convection number = $(1/X-1)^{0.8} (\rho_{v/} \rho_l)^{0.5}$
COP - Coefficient of Performance
Cp - Specific heat (J/kg.K)
d - Diameter of tube (m)
E - Enhancement factor
F( ) - General function
F_{fl} - Fluid-surface parameter
Fr - Froude number = $G^2/\rho_l^2.g.d$
G - Mass flux ($kg/m^2.s$)
g - Acceleration due to gravity (m/s^2)
h - Heat transfer coefficient ($W/m^2.K$)
h_{fg} - Latent heat of vaporization (J/kg)
k - Thermal conductivity (W/m.K)
L - Length of tube (m)
M - Molecular weight
Nu - Nusselt number
Pr - Prandtl number
p - Pressure
p_{sat} - Saturation pressure (N/m^2)
Q - Heat flux (W/m^2)
q - Specific heat flux ($W/m^2.s$)
Re - Reynolds number = $G (1-X) d/\mu_l$
S - Suppression factor
U - Overall heat transfer coefficient ($W/m^2.K$)
X - Quality of refrigerant
X_{tt} - Martinelli's coefficient

Greek letters:

c - Critical
Δ - Increment
μ - Viscosity ($N.s/m^2$)
ρ - Density (kg/m^3)
σ - Surface tension (N/m)
ψ - Dimensionless parameter (shah)

Subscripts:

cb - convective boiling
CBD - convective boiling dominant
eq - equivalent
i - inner
l - liquid phase
lo - liquid only
nb - nucleate boiling
NBD - nucleate boiling dominant
o - outer
pb - pool boiling
r - reduced
sat - saturation
TP - two phase
v - vapour phase
w - wall

I. INTRODUCTION

Evaporation heat transfer of refrigerant occurs in many applications like refrigerator, air-conditioner and heat pumping system by phase change. So, for design of these evaporators, proper correlations must be selected to calculate the boiling side of heat transfer coefficient for the design of evaporators. During the past few decades large number of correlations has been proposed for flow boiling of refrigerants for different types of tubes with different orientations.

Due to quick destruction of the ozone layer in the earth atmosphere has been noted which has been primarily related to the wide use of the CFC refrigerants which have been employed as the working fluids in many refrigeration, air conditioning and heat pump systems. Under the mandate of the Montreal Protocol the use of CFCs had been phased out and the use of HCFCs will also be phased out in a short period of time. Therefore we have to replace the CFCs by new alternative refrigerants In order to properly use these new refrigerants. We need to know their thermodynamic, thermophysical, flow and heat transfer properties. We realize that a much more detailed understanding of the flow boiling heat transfer of new refrigerants (R134a) is very important in the design of evaporators used in many current refrigeration and air conditioning systems.

In domestic refrigerators, HCs are successfully used because the charge of HC refrigerant has been reduced up to sufficiently small amount, which is around 50g, to avoid explosion caused by its leakage. Drop-in type systems were, therefore, able to be developed. In case of a room air conditioner, more than about 500g of hydrocarbon is required in a drop-in type system. If a small charged refrigeration system was applicable to the air conditioners, the refrigeration system using hydrocarbons could be widely accepted. Liquid stored in the condenser is greater part of the whole of refrigerant in the system. Therefore, the condenser and condensation would be one of the key technologies in order to apply the hydrocarbon to the air conditioners and other refrigeration and heat pump systems.

When refrigerant is evaporated in evaporator, different hydrodynamic conditions are encountered. This paper includes different correlations to predict heat transfer coefficient for flow boiling. The heat transfer coefficient can increase with quality, remain constant, or decrease with quality, depending on two parameters i.e. Boiling number, $Bo = q/(G \cdot h_{fg})$, and liquid to vapour density ratio (ρ_l/ρ_g), where G is mass flux, q is heat flux, and h_{fg} is latent heat of vaporization.

II. Boiling Correlation for Refrigerants:

During the past few decades, many aspects of boiling heat transfer have been investigated and a large number of correlations have been proposed for flow boiling of refrigerants for different geometry of tubes.

2.1. Boiling in horizontal smooth tubes:

The heat transfer coefficients of the evaporator, in this study, can be evaluated from conduction and forced convection in heat transfer theory as given below,

$$\frac{1}{h_r * A_i} = \frac{1}{U_o * A_o} - \frac{1}{h_w * A_o} - \frac{\ln\left(\frac{r_o}{r_i}\right)}{2*\pi*k_l} \quad \ldots (1)$$

Where h_r *and* h_w represents the refrigerant and water heat transfer coefficient ($W/m^{2o}C$) resp., Uo is the overall heat transfer coefficient ($W/m^{2o}C$), $\ln (r_o/r_i)/2\pi kL$ denotes wall resistance.

2.1.1. Sieder and Tate correlation:

Modified correlation for boiling in horizontal tubes with assumption that the fluid flow is fully developed and closed to that in the straight line.

$$h_W = 0.027*(\mathrm{Re})^{0.8}*(\mathrm{Pr})^{0.333}*k*\left[\frac{d_{o,in}^{\ 2}}{d_{i,out}^{\ 2} - d_{o,in}^{\ 2}}\right] \quad \ldots \quad (2)$$

The boiling heat transfer coefficient and pressure drop data have been taken. The present study, correlation equations of boiling heat transfer coefficients for R-22 in smooth tubes are

$$h_{TP} = h_{lo} * a * \left(\frac{1}{X_{tt}}\right)^b \quad \ldots (3)$$

$$X_{tt} = \left[\frac{1-X}{X}\right]^{0.9} * \left(\frac{\rho_v}{\rho_l}\right)^{0.5} * \left(\frac{\mu_l}{\mu_v}\right)^{0.1} \quad \ldots (4)$$

Where constant a=0.020 and b=2.00 for smooth tubes and a=0.023 and b=1.99 for grooved tubes.

$$h_{lo} = 0.023 * \left(\frac{G * d_i}{\mu_l}\right)^{0.8} * (\mathrm{Pr}_l)^{0.4} * \left(\frac{k_l}{d_i}\right) \quad \ldots (5)$$

Seider-Tate applies to "normal" fluids in turbulent flow in long, straight pipes, so:

$$0.7 < \mathrm{Prl} \le 160\,,\ \mathrm{Rel} \ge 10{,}000\ \&\ \frac{\mathrm{L}}{\mathrm{d_i}} \ge 60 \quad \ldots (6)$$

2.1.2. Lavin and Young correlation:

The heat transfer coefficients are experimentally measured at temperatures between −8 and 5 °C for Refrigerant R-134a.

$$h = 6.59 * h_l * \left(\frac{1+X}{1-X}\right)^{1.16} * \left(\frac{q}{h_{fg} * G}\right)^{0.1} \quad \ldots (7)$$

2.1.3. Shah correlation:

$$\frac{h_{TP}}{h_l} = \psi \quad \ldots (8)$$

Ψ can be determined from the largest value of Ψ_{nb} or Ψ_{cb}.

For horizontal tubes with $Fr_l \le 0.04$,

$$N = 0.38 * (Fr_l)^{-0.3} \quad \ldots (9)$$

For N>1.0,

$$\psi_{nb} = 1 + 46 * B_o^{0.5} \text{ For } B_o < 0.3*10^{-4} \quad \text{... (10)}$$

$$\psi_{nb} = 230 B_o^{0.5} \quad \text{For } B_o < 0.3*10^{-4} \quad \text{… (11)}$$

$$\psi_{cb} = \frac{1.8}{N^{0.8}} \quad \text{... (12)}$$

Mean deviation of theoretical correlation and experimental data for Shah Correlation is 14.0 to 17.2%.

2.1.4. Kandlikar correlation:

$$\frac{h_{TP}}{h_l} = C1 * CO^{C2} * (25 * Fr_l)^{C5} + C3 * B_o^{C4} * F_{fl} \quad \text{… (13)}$$

$$CO = \left(\frac{1-X}{X}\right)^{0.8} * \left(\frac{\rho_v}{\rho_l}\right)^{0.5} \quad \text{… (14)}$$

Condition	C1	C2	C3	C4	C5
Co < 0.5	1.1360	-0.9	667.2	0.7	0.3
Co > 0.5	0.668	-0.02	1058	0.7	0.3

In the above correlations the following definitions are used:
Dittus-Boelter correlation for single phase transfer:

$$h_l = 0.023 * \left(\frac{G * D_i * (1-X)}{\mu_l}\right)^{0.8} * (\mathrm{Pr}_l)^{0.4} * \left(\frac{k_l}{D_i}\right) \quad \text{… (15)}$$

$$h_{lo} = 0.023 * \left(\frac{G * D_i}{\mu_l}\right)^{0.8} * (\mathrm{Pr}_l)^{0.4} * \left(\frac{k_l}{D_i}\right) \quad \text{.. (16)}$$

Mean deviation of theoretical correlation and experimental data for Shah Correlation is 18.8 %.

2.2. Boiling in Vertical smooth tubes:

Gungor and winterton give the correlation for boiling in vertical tubes for two phase R 22 and R134a refrigerant.

$$h_{TP} = E * h_l + S * h_{pool} \quad \text{… (17)}$$

Where,

$$E = 1 + 24000 * B_o^{1.16} + 1.37 \left(\frac{1}{X_{tt}}\right)^{0.8} \quad \text{... (18)}$$

$$S = \left[\frac{1}{1 + \left(1.16 * 10^{-6} * E^2 * (\mathrm{Re}_l)^{1.17}\right)}\right] \quad \text{... (19)}$$

$$h_l = 0.023 * (\mathrm{Re}_l)^{0.8} * (\mathrm{Pr}_l)^{0.4} * \left(\frac{k_l}{d}\right) \quad \text{... (20)}$$

2.3. Boiling in fin tubes:

Generally saturated flow boiling in smooth tubes expressed by the combination of nucleate boiling and forced convection. The nucleate boiling is characterized by the presence of active nucleation sites, and it is expressed as a function of Boiling number, Bo.

$$B_o = \frac{q}{\Delta h * G} \quad \text{…. (21)}$$

The forced convection heat transfer is governed by mass flow rate and vapour quality, which is expressed as a function of Martinelli parameter, Xtt.

$$X_{tt} = \left[\frac{1-X}{X}\right]^{0.9} * \left(\frac{\rho_v}{\rho_l}\right)^{0.5} * \left(\frac{\mu_l}{\mu_v}\right)^{0.1} \quad \text{... (22)}$$

The present correlation for micro-fin tubes basically has the form of a smooth tube correlation. Generally, the ratio of flow boiling heat transfer coefficient, h_{tp}, in smooth tubes to single-phase convection heat transfer coefficient, h_l, can be represented as a function of Bo and Xtt.

$$\frac{h_{tp}}{h_l} = f(B_o, X_{tt}) \quad \text{... (23)}$$

Where,

$$h_l = 0.023 * (\mathrm{Re}_l)^{0.8} * (\mathrm{Pr}_l)^{0.4} * \left(\frac{k_l}{D_i}\right) \quad \text{… (24)}$$

Murata and Hashizume reported that the correlations of nucleate boiling heat transfer for micro-fin and smooth tubes, which are given in Eqs. (5) and (6).

$$h_{(PB,plain)} = 31.4 * \left[\frac{(Pc)^{0.2} * F_p}{M^{0.1} * (Tc)^{0.9}}\right] * (q_{PB})^{0.8} \quad \text{…(25)}$$

And,

$$h_{(PB,microfin)} = 40 * \left[\frac{(Pc)^{0.2} * F_p}{M^{0.1} * (Tc)^{0.9}}\right] * (q_{PB})^{0.8} \quad \text{... (26)}$$

For two-phase flow, the changes of liquid film thickness, velocity and void fraction are obviously dependent on quality, x. Although the trends of those changes are different between a smooth and micro-fin tube, these phenomena can be well implemented in terms of Xtt by introducing additional correction factors for micro-fin tubes. In addition to Bo and Xtt, the following operating parameters are considered in the present correlation (Table.1):

turbulence effects (enhancement factor-convective boiling)	$\frac{Gf}{\mu_l} * \mathrm{Re}_l$
fin height	δ/f
Evap. Temp.& surface tension	$\frac{P_{sat} * D_i}{\sigma}$
fluid properties	P_r

The turbulence effects generated by micro-fins also produce considerable influence on nucleate boiling heat transfer. These two factors happen to simultaneously improve the evaporation heat transfer at low vapour quality.

Combining and rearranging the non-dimensional parameters, the present correlation for the evaporation heat transfer in micro-fin tubes is given by,

$$\frac{h_{tp}}{h_l}=\left[C1*(B_o)^{C2}*\left(\frac{P_{sat}*D_i}{\sigma}\right)^{C3}+C4*\left(\frac{1}{Xtt}\right)^{C5}*\left(\frac{Gf}{\mu_l}\right)^{C6}\right]$$
$$*(\mathrm{Re}_l)^{C7}*(\mathrm{Pr}_l)^{C8}*\left(\frac{\delta}{f}\right)^{C9} \quad \text{... (27)}$$

The coefficients in Eq. (7), which are given in Table 2, were determined by a multiple regression analysis with the database listed in Table 2.

Coefficient	*C1*	*C2*	*C3*	*C4*	*C5*
Value	*0.00962*	*0.1106*	*0.3814*	*7.6850*	*0.51*
Coefficient	*C6*	*C7*	*C8*	*C9*	
Value	*0.7360*	*0.2045*	*0.7452*	*0.1302*	

Few updated correlations which let plain finned horizontal tube heat exchangers be modelled with a high degree of accuracy. In the existing literature, many correlations for characterizing boiling inside plain tubes can be found. Some have been studied by the authors, and most of them belong to one of the next three categories:

- Superposition methods (i.e. Chen correlation).
- Enhancement models (i.e. Shah Correlation).
- Asymptotic models (i.e. Steiner and Taborek correlation).

Superposition and asymptotic models consider two thermal mechanisms; nucleate boiling and convective boiling which are included by means of expressions such as

$$h_{tp}=\left[(h_{nb})^n+(h_{cb})^n\right]^{\frac{1}{n}} \quad \text{… (28)}$$

Where n= 1 for Chen correlation, and n= 3 for Steiner and Taborek's correlation. Asymptotic methods have been proven to be very suitable for the characterization of flow boiling, since they cover a wide range of refrigerants and take into account most of the regimes inside the pipes. Nucleate pool boiling correlations are frequently used to account for the nucleate boiling contribution to the heat transfer coefficient. These are based on,

- Reduced pressure (i.e. Cooper)
- Fluid specific (i.e. Gorenflo correlation)

2.3.1. Cooper correlation:

The correlation proposed by Cooper for the heat transfer coefficient for nucleate pool boiling is a function of the reduced pressure and includes the effect of wall roughness and heat flux. The applicability of this expression in terms of reduced pressure, pr, and molecular weight, M, is $0.001\leq \mathrm{Pr}\leq 0.9$ and $2\leq M\leq 200$

$$h_{pb}=55*(\mathrm{Pr})^{0.12}*(\log_{10}\mathrm{Pr})^{-0.55}*M^{-0.5}*q^{0.67} \quad \text{… (29)}$$

2.3.2. Gorenflo correlation:

It considers a separate treatment of the main groups of variables: properties of the fluids, nature of the heated surfaces, operating parameters, etc.

$$\frac{h}{h_0}=C_W*F(\mathrm{Pr})*\left(\frac{q}{q_0}\right)^{n(\mathrm{Pr})} \quad \text{… (30)}$$

Where C_W stands for the effect of the properties of the heated surface, which may be obtained by the following expression:

$$C_W=\left(\frac{Ra}{Rao}\right)^{0.133} \quad \text{... (31)}$$

With R_{a0} = 0.4μm being the roughness parameter reference value.

The reference value a0 is calculated by using the Stephan and Preusser correlation with a reduced pressure of P_r=0.1 and a heat flux of q_o=20,000 W/m^2. This correlation is given by,

$$\frac{h}{ho}=Fq*Fp*Fwr*Fwm \quad \text{… (32)}$$

Where, Eq is a function of the heat flux ratio between the flux at the wall and a reference heat flux; Fp is a function of the reduced pressure; Fwr takes into account the micro-structure; Fwm considers the material properties.

$$h=\frac{k_l}{d_o}\left[\frac{q_o*d_o}{Ts*k_l}\right]^{0.674}*\left(\frac{\rho_v}{\rho_l}\right)^{0.156}*\left[\frac{\Delta h_l*(d_o)^2}{h_l^{\,2}}\right]$$
$$*\left[\frac{(h_l)^2*\rho_l}{\sigma*d_o}\right]^{0.350}*\left(\frac{\mu_l*Cp_l}{k_l}\right) \quad \text{.... (33)}$$

As this equation is only suitable for a reduced pressure of Pr= 0.3, h_0 is obtained in two steps. The bubble diameter at departure d_0 is given by,

$$d_0 = 0.0149 * \beta * \left[\frac{2 * \sigma}{g(\rho_l - \rho_v)} \right]^{0.5} \quad \text{.... (34)}$$

Where the contact angle β is 35^0.

The Gorenflo correlation is valid only for values of the reduced pressure $0.1 \leq Pr \leq 0.9$.

2.4. Boiling in Minichannels:

The flow boiling correlation by Kandlikar (1990, 1991) as reported in Kandlikar et al. (1999) is as follows for boiling in channel:

h_{TP} = larger of $h_{TP,\,NBD}$ or $h_{TP,\,CBD}$

Where $h_{TP,\,NBD}$ and $h_{TP,\,CBD}$ are two-phase heat transfer coefficients in the nucleate boiling dominant and convective boiling dominant regions, as given by the following equations.

$$h_{TP,NBD} = \left[0.6683 * Co^{-0.2} * (1-X)^{0.8} * f(Fr_{lo}) * h_{lo}\right] + \left[1058.0 * B_o^{\,0.7} * (1-x)^{0.8} * F_{fl} * h_{lo}\right] \quad \text{… (35)}$$

$$h_{TP,CBD} = \left[1.136 * Co^{-0.9} * (1-X)^{0.8} * f(Fr_{lo}) * h_{lo}\right] + \left[667.2 * B_o^{\,0.7} * (1-x)^{0.8} * F_{fl} * h_{lo}\right] \quad \text{… (36)}$$

Where,

$$CO = \left(\frac{1-X}{X}\right)^{0.8} * \left(\frac{\rho_v}{\rho_l}\right)^{0.5} \text{ And} \quad \text{... (37)}$$

$$B_o = \frac{q}{h_{fg} * G} \quad \text{… (38)}$$

x is the quality and F_{fl} is the fluid-surface parameter.

Refrigerant	F_{fl}
R22	2.20
R134a	1.63

The single phase heat transfer coefficient is given by the Gnielinski correlation.

$$h_{lo} = \left[\frac{\mathrm{Re}_{lo} * \mathrm{Pr}_l * \left(\frac{f}{2}\right) * \left(\frac{k_l}{d}\right)}{1 + 12.7\left[(\mathrm{Pr}_l)^{\frac{2}{3}} - 1\right] * \left(\frac{f}{2}\right)^{0.5}} \right] \quad \text{… (39)}$$

For $10^4 \leq Re_{LO} \leq 5.10^6$ and,

$$h_{lo} = \left[\frac{(\mathrm{Re}_{lo} - 1000) * \mathrm{Pr}_l * \left(\frac{f}{2}\right) * \left(\frac{k_l}{d}\right)}{1 + 12.7\left[(\mathrm{Pr}_l)^{\frac{2}{3}} - 1\right] * \left(\frac{f}{2}\right)^{0.5}} \right] \quad \text{… (40)}$$

For $2300 \leq Re_{LO} \leq 5.10^4$

Where friction factor,

$$f = \left[1.58 \ln(\mathrm{Re}_{lo}) - 3.28\right]^{-2} \quad \text{… (41)}$$

III. Conclusion

Proper estimation of boiling heat transfer coefficient is very essential for optimum sizing of evaporators of various applications working on VCR cycle. Optimum sizing calls for minimum pressure drop and hence requires minimum power consumption for given duty. A review of various commonly referred correlations for boiling heat transfer is presented in the paper. Correlations for horizontal and vertical tube orientations of smooth, rough, finned and grooved surface have been reviewed. This paper is useful to choose simple and efficient form of generalize correlation of flow boiling heat transfer for refrigerants R22 and R134a.

REFERENCES:

i. **G.A. Longo and A. Gasparella**, Refrigerant R134a vaporization heat transfer and pressure drop inside a small brazed plate heat exchanger, International Journal of Refrigeration 30 (2007), pp. 821-830.

ii. **J.R. Garcıa-Cascales**, F. Vera-Garcıa, J.M. Corberan-Salvador & David Fuentes, Assessment of boiling heat transfer correlations in the modelling of fin and tube heat exchangers, International Journal of Refrigeration 30 (2007), pp. 1004-1017.

iii. **Rin Yun, Yongchan Kim, Kookjeong Seo & Ho Young Kim**, A generalized correlation for evaporation heat transfer of refrigerants in micro-fin tubes, International Journal of Heat and Mass Transfer 45(2003), pp. 2003–2010.

iv. **Satish G. Kandlikar and Mark E. Steinke,** Flow Boiling Heat Transfer Coefficient In Minichannels – Correlation and Trends, Mechanical Engineering Department, Rochester Institute of Technology, Rochester, New York 14623, USA

v. **S.G.Kandlikar**, a general correlation for saturated two phase flow boiling heat transfer inside horizontal and vertical tubes, Mechanical Engineering Department, Rochester Institute of Technology, Rochester, New York 14623, USA

vi. **Uthen Kuntha & Tanongkiat Kiatsiriroat**, Boiling heat transfer coefficient of R-22 refrigerant and its alternatives in horizontal tube: small refrigerator scale, Songklanakarin J. Sci. Technol., 2002, pp. 243-253.

CFD Analysis of an Axial Flow Pump

Tushar Chourushi,
Asst. Prof., Mechanical Engineering Dept.,
Gandhinagar Institute Of Technology

***Abstract*—CFD (computational fluid dynamics) analysis generally gives us an idea about the solution of the given problem. In this paper an Axial Flow Pump was taken for the study. Appropriate meshing techniques and boundary conditions were implemented on to the specimen geometry. The results were then obtained and conclusion was made out of it.**

***Index Terms*—Ansys ICEM-CFD, Ansys CFX**

I. INTRODUCTION

The paper mainly focuses on to the analysis of the flow through an axial flow pump, with water as a medium. The analysis is done by with full the geometry using the Computational fluid dynamics. Analysis is done assuming the flow to be steady and interface is given for mesh.
The purpose of the project is to analyze the flow through an Axial Flow Pump with different mass flow rates, and obtain the optimized mass flow rate by comparing the plots.

- Geometry is created using the ANSYS ICEMCFD software tool.
- Meshing too was done using ANSYS ICEMCFD tool.
- Meshes are made smoothen up to a limit which is acceptable in ANSYS CFX.
- Pre-processing, Solving and Post-processing were done with the use of ANSYS CFX tool.
- Results were compared and graphs were plotted as required.

II. GOVERNING EQUATIONS

One of the most useful concepts in the performance evaluation of turbo-machinery is the control volume approach as applied to the basic laws of conservation of mass, momentum, and energy.

The generalized equation relating the laws of mechanics (system approach) to the fluid which is instantaneously contained within the control volume is given by

$$\frac{dN}{dt} = \iint n\rho V \cdot dA + \frac{\partial}{\partial t}\iiint n\rho dv \qquad (2.1)$$

Where, N is any extensive property such as mass, linear momentum, angular momentum, or stored energy (internal plus kinetic energy), and n is the extensive property per unit mass. The equations for continuity, linear momentum, angular momentum, and energy based on these equations are given by substituting n = N/ M = 1, V, r × V, e, respectively, in above Equation 2.1.

$$0 = \iint_{c.s.} \rho V \cdot dA + \frac{\partial}{\partial t}\iiint_{c.v.} \rho dv \qquad (2.2)$$

$$\iint_{c.s.} F dA + \iiint_{c.v.} B\rho dv = \iint_{c.s.} V(\rho V \cdot dA) + \frac{\partial}{\partial t}\iiint_{c.v.} V(\rho dv) \qquad (2.3)$$

$$\iint_{c.s.} r \times F dA + \iiint_{c.v.} r \times B\rho dv = \iint_{c.s.} (r \times V)(\rho V \cdot dA) + \frac{\partial}{\partial t}\iiint_{c.v.} (r \times V)\rho dv \qquad (2.4)$$

$$\frac{dq}{dt} - \frac{dW_s}{dt} - \iiint_{c.v.} B \cdot V \rho dv = \iint_{c.s.} h_o \rho V \cdot dA + \frac{\partial}{\partial t}\iiint_{c.v.} e_o \rho dv \qquad (2.5)$$

Here F is the surface (normal and tangential) force per unit area acting on the control surface; B is the body force per unit mass, such as gravity, acting inside the control volume; r is the distance from the origin of the coordinate system; q is the heat transferred to the control volume; Ws is the work done on the control volume by the rotor and by shear forces; and ho is the stagnation enthalpy.

III. GENERAL METHOD FOR MESH CONNECTION

The General Connection interface model is a powerful way to connect regions together. The General Connection option is necessary when the frame of reference or pitch changes across the interface

This model produces a steady state solution to the multiple frame of reference problems, with some account of the interaction between the two frames. The quasi-steady approximation involved becomes small when the through flow speed is large relative to the machine speed at the interface. Frozen Rotor analysis is most useful when the circumferential variation of the flow is large relative to the component pitch. This model requires the least amount of computational effort of the three frames change/mixing models.

IV. Problem Description

An Axial Pump is divided into two main sections; first section is termed as rotor (or impeller) and the second section is called as stator (or diffuser). For the Full channel analysis full impeller and diffuser geometry is made.

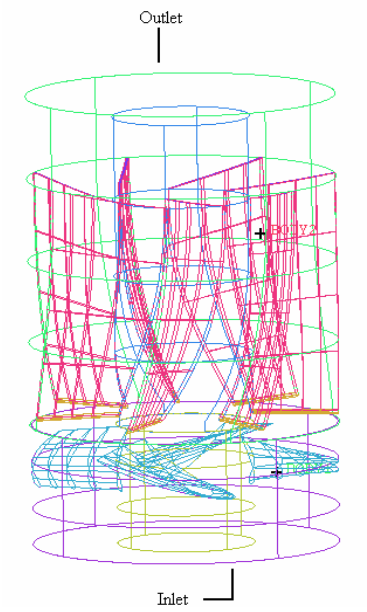


Fig: 1. Full geometry: both impeller and diffuser

The geometry is made in ANSYS ICEMCFD. The above is the Propeller 35 data geometry. The term 35 is the designation for the staggered angle used in the impeller. The blades of impeller and diffuser does maintain the pitch ratio, it is advised for better pump performance. The diffuser blades do make a passage for the flow, which enable to enhance the pressure head.

Assumptions made:

- Steady State
- Incompressible Fluid
- Subsonic regime
- Conservative
- No – slip boundary conditions at the walls

In this case Shear Stress turbulence model is used, $k-\omega$ model are slightly better than k – e, but still does not behave well. More modern variants like Menter's SST k-omega model behave better. So, in our case we took SST $k-\omega$ model.

V. Geometry Creation

The Geometry creation was done in ANSYS ICEMCFD. The propeller 35 data is used to create an Axial Flow Pump; the blade is set to a staggard angle of 35. The rectilinear co-ordinate system of Geometry creation was employed.

Blade geometry is created with the Lofting Surface approach over several curves. This is done to get the smooth surface.

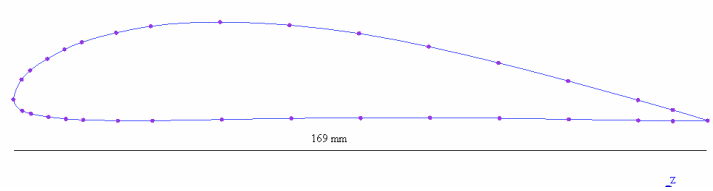


Fig: 2. The blade profile used in impeller is as shown above.

The geometry of the impeller prop-35 with 4 blades is made using Ansys-ICEM-CFD tool. The impeller hub carries the four blades, with shroud acting as a casing. No clearance is provided in the given model used. The Hub diameter of 238 mm, and shroud diameter of 530 mm is used. In case of the Diffuser 7 blades are used. These numbering depends upon the best performance of the turbine for the above case.

VII. Mesh Generation

Unstructured meshes are used for more complex and odd geometries where a structured mesh is difficult to create.
It is difficult to define, a priori, the mesh size. The required mesh size depends on the purpose of the simulation.
A good low-Re mesh with resolved boundary layers typically has around 1,000,000 cells. For more accurate simulations with resolved boundary layers the mesh should have a y+ for the first cell which is below 10 for turbo-machinery applications. In present case y+ is maintained below 12.

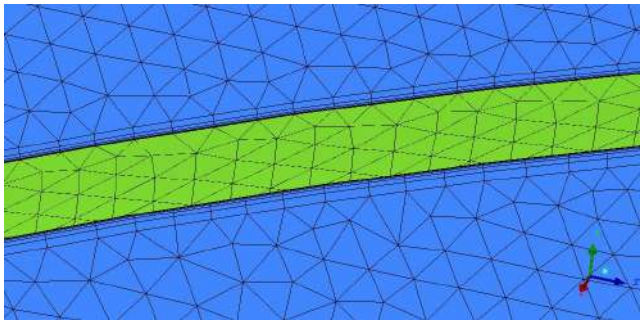


Fig: 3. Prism layers onto the blades

Total 6 prism layers are used on to the blade. This numbering is dependent onto the flow velocity and Reynolds number. It also depends upon the type of fluid used.

After meshing, mesh files are imported into the ANSYS CFX-Pre. With the help of this software tool the boundary conditions and initial conditions are given for the further calculation.

VIII. Boundary Conditions

In our case, Dirichlet boundary condition is suited best for both inlet and outlet. At walls No-slip boundary condition is used.

- For the inlet, mass flow rate of water entering is 250 kg/s.
- For the outlet, static pressure outlet.
- At the walls, No slip condition is employed.
- Domain interface boundary is considered at the impeller and diffuser interface.

IX. Solving and Simulations

After the boundary conditions and initial condition is set, the solver is set to run for 300-500 iterations. The iteration can be edited as per the solution. The convergence limit for all the fluid flow equations is set up to 10^{-5} which is the best approximation for the turbo machinery.

Simulation is later done with the result obtained by solving the conservation equation in the CFX solver.

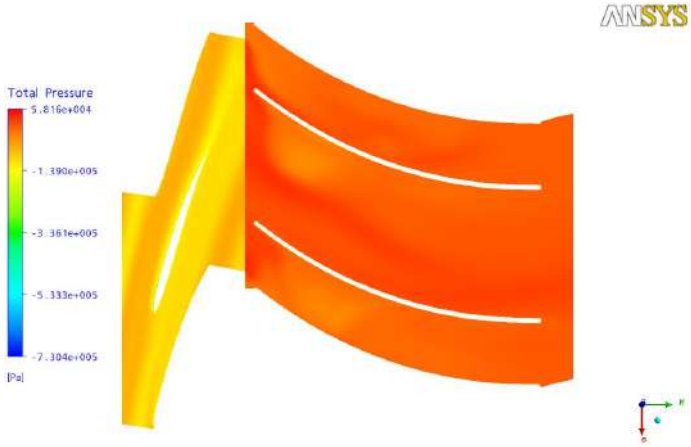


Fig: 4. Total Pressure Head distribution of Blade to Blade profile of Pump

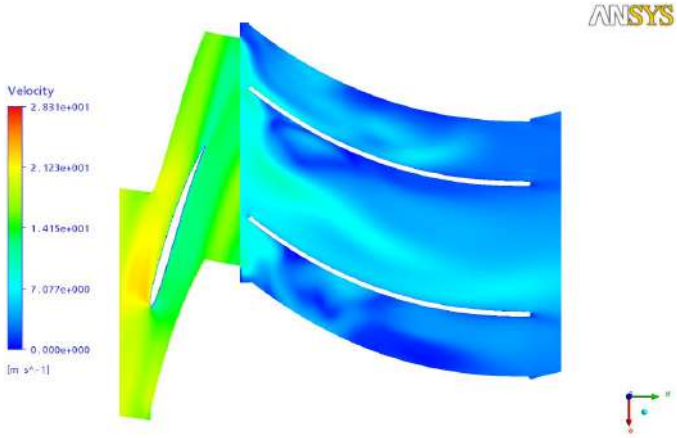


Fig: 5. Velocity variation of Blade to Blade profile of Pump

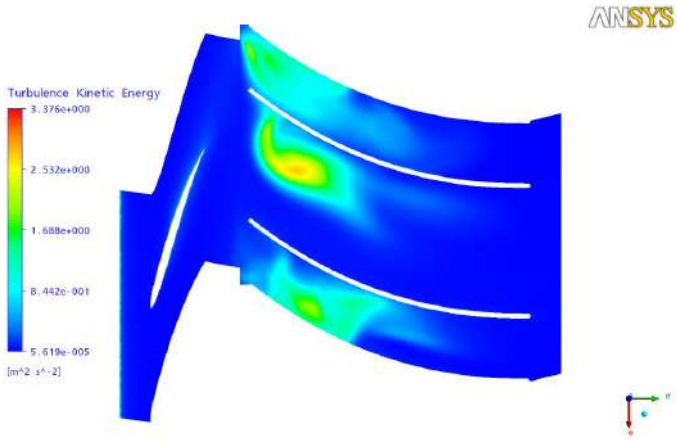


Fig: 6. Turbulence Kinetic Energy of Blade to Blade profile of Pump

X. Results and Conclusion

In solver once the solution achieved convergence, it is said to be approached to a conclusion. After this we have to check the fluxes (Mass Flow Rate) between inlet & outlet and the static pressure difference between inlet and outlet. This Numerically solved result is then validated

with the actual Flow field result; in our case “Flow parameters” gives confirmation about the correctness of the result.

The flow is considered as being incompressible and turbulent. The turbulence model SST (Shear Stress Turbulence, model) has been used for validation process with convergence set at 10^{-5}. This convergence criterion is also utilized for validation of the pressure and velocity fields.

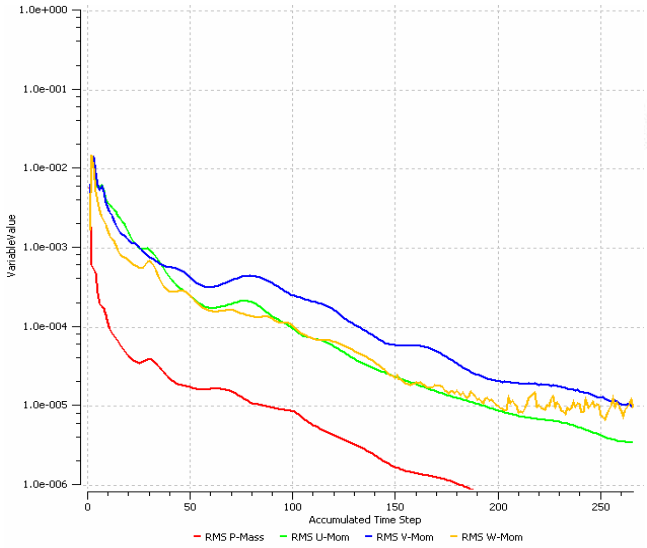


Fig: 7. Residual Convergence Plot for Momentum and Pressure

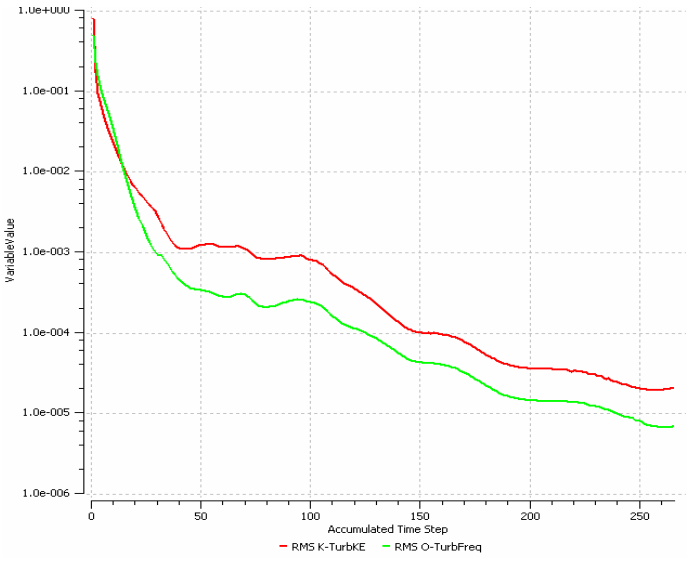


Fig: 8. Residual Convergence Plot for Turbulence KE & Omega Frequency

Analysis Result of an Axial Flow Pump:

In our paper three cases different mass flow rates were considered:

- 250 kg/s
- 275 kg/s
- 310 kg/s

Performance Curve Graphs:

Efficiency:

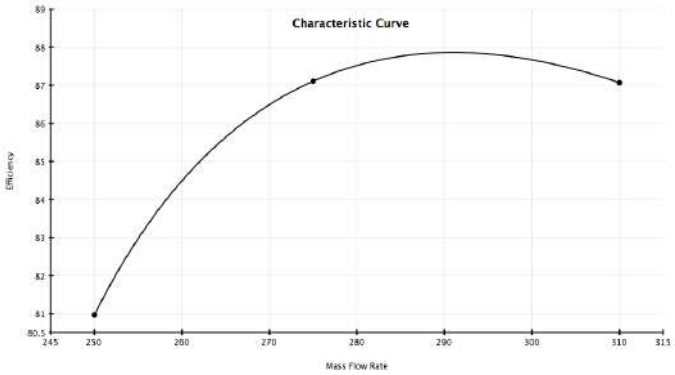


Fig: 9. Efficiency Curve

Pressure Head:

The Figure below shows “Total Head Pressure” developed by the Pump considering all loses occurred when Pump is running with different mass flow rates.

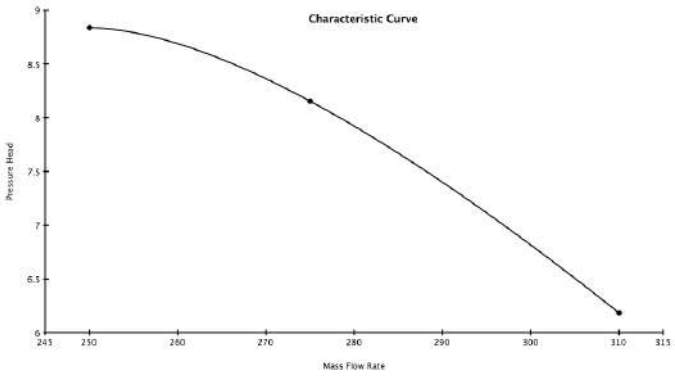


Fig: 10. Pressure Head Curve

Conclusion:

From the above graphs obtained, it is to be noted that the Performance curve is very important criteria for determining the results. It is to be noted that the efficiency graph shows that at 275 kg/s mass flow rate, the efficiency of the Axial Pump is higher compared to both of the other mass flow rates. So, it shows that the mass flow rate 275 kg/s is an optimum design for operating the Axial Pump of the given dimension. Though, there is a decrease in Total pressure Head at this mass flow rate. So, the decision has to be made as to which criterion is important for the user.

XI. Recommendation Of Future Work

- The presented methodology establishes a range for optimum selection of an Axial Pump with high efficiency, showing to be a good indication of the determination of an initial geometry of this turbo-machine type.
- Different Advection schemes can be employed Upwind, High resolution etc.
- Transient Simulation type can be used instead of Steady.
- Instead of SST Turbulence model some other Turbulence model can be studied.
- Different General connections can be used for Mesh connections.
- Other influences must be considered for the improvement of this design methodology such as secondary flows, tip clearance, which were not considered here.
- Also cavitations effect may be analyzed in more detail.

References

1) *Fluid Dynamics and Heat Transfer of Turbomachinery, by Budugur Laxminarayana*

2) *Journal of the Brazilian Society of Mechanical Sciences and Engineering, technical Papers*

3) *J Blazek, Elsevier. "Computational fluid dynamics: principles and applications"*

4) *"Computational Fluid Dynamics Volume" – III Fourth Edition, Klaus A. Hoffmann & Steve T. Chiang*

5) *"Computational Fluid Dynamics The Basics with Applications", John D. Anderson Jr.*

6) *"Introduction to turbulence modeling", Goodarz ahmadi, Department of mechanical and aeronautical engineering, Clarkson university, Potsdam, ny 13699-5725*

7) *"Notes on Turbo-machinery", Dale B. Andrews – Editor, ©2005 Lawrence Pumps Inc. 371 Market St., Lawrence. MA. 01843*

8) *"Fluid Mechanics, Thermodynamics of Turbomachinery", S.L. Dixon, B.Eng., PH.D.*
9) *Senior Fellow at the University of Liverpool*

10) *"Practical Programming in Tcl/Tk", Fourth Edition, By Brent B. Welch, Ken Jones, Jeffrey Hobbs*

11) *ANSYS ICEM CFD 11.0 tutorial manual*

12) *ANSYS CFX, Release 11.0 tutorial notes*

13) *AGARD 1989, "Blading design for Axial Flow Turbomachines" AGARD LS 167*

14) *Ainley, D.G.,1948, "Performance of Axial Flow Turbines", Proc. Inst Mech Eng. (London), Vol. 159, p. 230 (1951).*

15) *Anderson, O. 1.., 1987, "Calculation of Three-Dimensional Boundary Layers on Rotating Turbine Blades, J. Fluids Eng., Vol. 109, pp. 41-50.*

16) *Balsa, T. F. and Mellor, G. 1.., 1972, "The Simulation of Axial Compressor Performance Using an Annulus Wall Boundary Layer Theory", AGARD AG 164, p. 373*

17) *Bindon, J. P., 1987, "Pressure and Flow Field Measurements of Axial Flow Turbine Tip Clearance Flow in a Cascade," Proc. Inst. Mech. Eng. (London), pp. 43-52; also see ASME Paper 87-GT-230 and J. Turbomachinery, Vol. 115, pp. 257-263, 1989.*

18) *Coakley, T. J., 1983, "Turbulence Modeling Methods for the Compressible NavierStokes Equations," AIAA Paper 83-1693.*

19) *http://www.cfd-online.com/Wiki/Best_practice_guidelines_for_turbomachinery_CFD*

20) *http://www.grc.nasa.gov/WWW/RTT/*

21) *http://en.wikipedia.org/wiki/Pump*

Design, Improvement and Thrust Bearing Analysis of Oil Expeller Machine

Mehul.K.Modh[1*]
M.Tech CAD/CAM Student,
U. V. Patel College of Engineering,
Kherva, Mehasana.,Gujarat, India
[*]Corrosponding Author
modh_mehul@yahoo.co.in

J.R.Mevada[2]
[2] Asst. Professors, U. V. Patel College of Engineering,
Kherva, Mehasana.,Gujarat, India

Abstract There is wide variety of food oils available in market such as Sun flower oil, Soya bean, Rap-Mustard seeds oil, Palm Kernal oil, Coconut oil, Ground nut oil, Castor oil, Sesame oil, cotton seeds oil etc. To manufacture all above oil products variety of oil expeller machines required. Among these, Screw type oil Expellers (presses) are advanced oil processing machinery, characterized by their high oil output rate with good quality, simple design, easy to use and continuous operation. They can use for various raw materials, such as peanut, beans, rape seeds, cotton seeds, sesame, sunflower seeds, copra etc.

Here the presented work describes design steps of some of the components of the oil expeller machine, Calculation of existing bearing's life, Suggest a new bearing with higher life rating and static analysis of races and ball in Pro-E and Ansys Workbench. Also the analytical work is done and compared with the software analysis. In addition the software results are also compared

Key Words: Oil Expeller Machine, Design, Thrust Ball Bearing, Ansys, Pro-E

1. Introduction

Screw type oil presses are advanced oil processing machinery, characterized by their high oil output rate with good quality, simple design, easy to use and continuous operation. They can use for various raw materials, such as peanut, beans, rape seeds, cotton seeds, sesame, sunflower seeds, copra etc.

Abbreviations

TPD	=	Tones per Day
N	=	Revolutions per Minute Main Shaft (RPM)
N_1	=	Revolutions per Minute Motor shaft (RPM)
N_{max}	=	Maximum rotational speed
P	=	Motor Power
p	=	Pressure
D_1	=	Motor Pulley Diameter
D_2	=	Gear Box Pulley Diameter
T	=	Main Shaft Torque
τ	=	Shear Stress
Fa	=	Axial or Thrust Force
M	=	Minimum axial load factor
Cdyn	=	Dynamic Capacity of Bearing

According to the online resources, Bamgboye and Adejumo [1] developed an expelling machine for extracting oil from decorticated sunflower seeds. The machine was tested at auger speeds of 30, 40, 50 rpm respectively and three throughputs. Results showed that performance efficiencies increased with auger speed and throughput. Expelling efficiency of over 70% was obtained

Alessandra Maria Sabelli [2] represented a first step into the development of an extraction tool that maximizes the extraction of oil from moringa seeds, and consequently the consumption of the seeds themselves, not exploited so far. The author designed the press and carried out an experiment to investigate the effect of pressure applied on seed vs. oil extracted

The uses of various investigational techniques were described and illustrated with examples of work by D. Scott of the National Engineering Laboratory. [3] Also methods of diagnosis of impending bearing failure and its prevention are outlined.

The main objective of this work is to design and improvement by replacing the existing bearing. Also carry out analytical and software analysis of thrust ball bearing used in machine.

2. Product Design

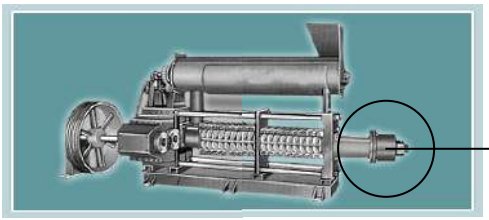


Fig.1. Double Chamber Oil Expeller Machine

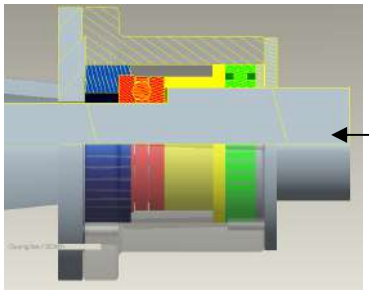


Fig.2. Location of Thrust Ball Bearing

2.1 Problem Description

1. Thrust ball bearing of m/c fails once in 2 to 3 months.
2. Type of failure It breaks into pieces

2.2 Technical Details

These details are taken from the machine catalogue

1. Steam rate 2 Kg/Cm2
2. Quantity to be crushed 25 to 30 TPD
3. Main Shaft rotation N = 22 RPM
4. Motor rotation N_1 = 940 RPM
5. Motor Capacity 85 HP
6. Thrust bearing No (ZKL Company) 51326 M
7. Pr. required expelling oil from mustard seeds 10 MPa to 15 MPa.
8. Motor Pulley Dia. D_1 = 240 mm
9. Gear Box Pulley Dia. D_2 = 1000 mm

2.3 Calculate Gear Ratio and Verify Ma in Shaft Rotation [4]

- Find input speed for Gear Box

  For belt drive

$$\frac{D_1}{D_2} = \frac{N_2}{N_1}$$

$$\frac{240}{1000} = \frac{N_2}{940}$$

N_2 = 225.6 RPM

- For Gear Drive

Gear Ratio $G_1 = \frac{T_3}{T_2} = \frac{D_3}{D_2}$

$$G_1 = \frac{54}{18} = \frac{48}{16} = 3$$

Similarly, $G_2 = \frac{T_5}{T_4} = \frac{63}{18} = 3.5$

Gear Ratio

$G = G_1 * G_2 = 3 * 3.5 = 10.5$

Main Shaft Rotation N = I/P speed at Gear Box * Gear Ratio G

N = 225.6 * 10.5

N = 21.48 RPM

Say N = 22 RPM

2.4 Main Shaft Design. [4]

1 HP = 0.754699 KW

Motor Power P = 85 HP

P = 64.15 KW

Power $P = \frac{2\pi NT}{60}$

$$64.15 * 10^3 = \frac{2\pi * 22 * T}{60}$$

$T = 27.84 * 10^3$N m = $27.84 * 10^6$N mm

Shear Stress for 45C8 Steel [20]

$\tau_u = 500 N/mm^2$ Say $\tau = 150\ N/mm^2$ Now

Torque $T = \frac{\pi * d^3 * \tau}{16} = \frac{\pi * d^3 * 150}{16} =$

$27.84 * 10^6$ N mm

Main Shaft Dia. d = 98.157 mm

Say d = 110 mm

2.5 Small Chamber Design [4]

$T = 27.84 * 10^6$N mm = F * r = F * 55

$F = 506.18 * 10^3$ N

$$Pr. = \frac{F}{A} = \frac{F}{\pi * d_h * L} = 140\ N/mm^2 = \frac{506.18 * 10^3 N}{\pi * 140 * L}$$

Where ,

d_h = Warm's core diameter

Pr. = Pressure required to expel oil from mustard seeds 10 MPa to 15 MPa.

Hence taking Pr. = 14 Mpa = 140 N / mm^2 (Q 1 MPa =10 N / mm^2)

Length of Chamber L = 822.4 mm

Take L = 850 mm

$$Pr. = \frac{F}{A} = \frac{F}{\pi * D * (D - d_h)} = 140\ N/mm^2 = \frac{506.18 * 10^3 N}{\pi * D * (D - 140)}$$

Chamber inner Dia. D = 147.79 mm

Take D = 170 mm

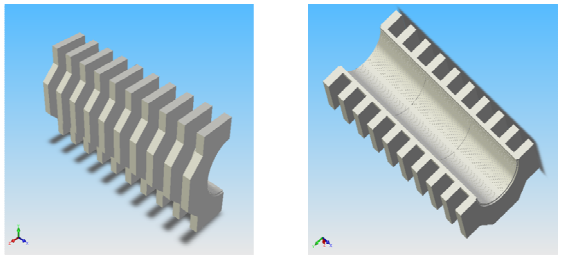


Fig.3. Small Chamber Casing

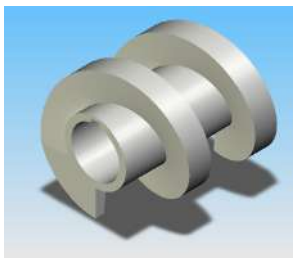


Fig.4. Warm inside the Chamber

3 Bearing Diagnosis

Table 1 Bearing Data [5]

Dimensions						Basic Load Rating		Fatique load limit	Limiting Speed for Lubrication with		Bearing Designation
d	D	d_1	D_1	H	r_s min	Dynamic C_a	Static C_{oa}	P_u	Grease	Oil	
mm						kN		kN	min^{-1}		
130	170	170	132	30	1.0	119.0	406.0	43.84	1400	1900	51126
	225	220	134	75	2.1	367.0	1070.0	109.85	750	1000	51326
	270	265	134	110	4.0	637.0	2010.0	199.10	560	750	51426

Abutment and Fillet Dimensions				Weight	Minimum Axial Load Factor
d	d_a min	D_a max	r_a max	~	
mm				kg	
130	154	146	1.0	1.870	0.650
	186	169	2.0	13.300	6.200
	213	187	3.0	32.000	18.000

3.1 Minimum Axial Load Required [5]

At higher rotational speeds danger of ball sliding between ring raceways can occur because of centrifugal forces, if axial load Fa drops under minimum value. Minimum value Fa is calculated from equation:

$$Fa\ min = M\ (N_{max} / 1000)^2\ [KN]$$

$$Fa\ min = 6.2\ (22 / 1000)^2 = 3*10^{-3}\ [KN]$$

Where, (Refer Table1)

Fa min – minimum axial load [KN]

N_{max} – maximum rotational speed [min^{-1}]

M – Minimum axial load factor

If the axial load is smaller than Fa min, or if bearing relieving comes into being during operation, e.g. of one ball row in double direction bearing, or of one bearing when using a pair of single direction thrust bearings, it is necessary to secure minimum load, e.g. by means of springs.

3.2 Calculate Thrust Load [6]

$$\text{Thrust Load } Fa = \frac{\pi * D^2 * P}{4*10000} KN = \frac{\pi * 170^2 * 100}{40000}$$

Fa = 226.865 KN

Calculate Nominal Thrust Bearing Life for Bearing No 51326M

Thrust Bearing Life,

$$L_{10} = \frac{10^6}{60 * 22} * (C_{dyn} / Fa)^{(10/3)}$$

$$L_{10} = \frac{10^6}{60 * 22} * (367 / 226.865)^{(10/3)}$$

L_{10} = 3765 hrs

L_{10} = 156.875 days

Lh_{10} = 5.2 Months

Here we get the life of bearing 5.2 months but actually when machine is operated 24 hours its designed life should be 50 to 60

Step 1

Select the bearing life of 3 years = 26280 hrs (See the Table No 2)

Thrust Bearing Life

$$L_{10} = \frac{10^6}{60 * 22} * (C_{dyn} / Fa)^{(10/3)}$$

$$26280 \text{ hrs} = \frac{10^6}{60 * 22} * (C_{dyn} / 226.865)^{(10/3)}$$

C_{dyn} = 657.39 KN

Step 2

Identify candidate bearings with required rating

Bearing no 51426 is selected

C_{dyn} = 637 KN (See table No 1)

Kh. So it is desirable to select a newer bearing with the desired life. (Refer Table 2)

3.3 Bearing Selection Process

- Calculate the dynamic load rating, C_{dyn} For the required bearing life.
- Identify candidate bearings with required rating
- Select bearing with most convenient geometry, also considering cost and availability

Table 2 Bearing-Life Recommendations Table [5]

Sr. No.	Type of Application	Life, Kh
1	Instruments and apparatus for infrequent use	Up to 0.5
2	Aircraft engines	0.5 – 2
3	Machines for short or intermittent operation where service interruption is of minor importance	4 – 8
4	Machines for intermittent service where reliable operation is of great importance	8 – 14
5	Machines for 8-h service	14 –

	which are not always fully utilized	20
6	Machines for 8-h service which are fully utilized	20 – 30
7	Machines for continuous 24-h service	50 – 60
	Machines for continuous 24-h service where reliability is of extreme importance	100 – 200

3.4 Bearing Analysis Result and Discussion

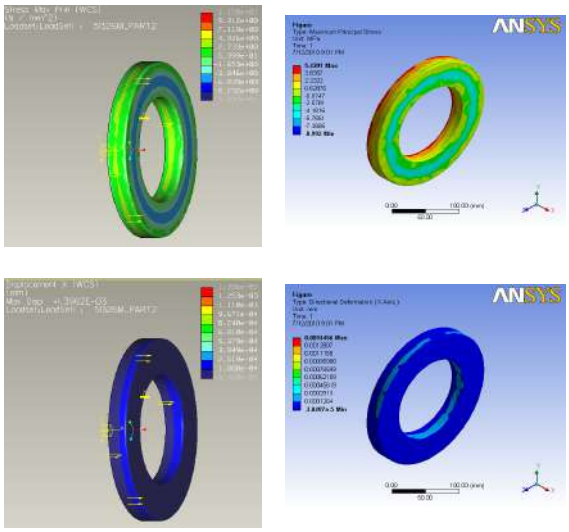


Fig.5. Analysis of Right side race in Pro e and Ansys work Bench (Selected Fig Only)

4 Analytical Solutions

4.1 Formulas for Right Side and Left Side Races [7]

Left Side Race Area= $\frac{\pi}{4}\left(d_1^{\,2} - d^2\right)$

Right Side Race Area= $\frac{\pi}{4}\left(D^2 - D_1^{\,2}\right)$

Principle Stress $\sigma = \frac{P}{A}$

Axial Deformation or Displacement

$\delta_l = \frac{P*l}{A*E}$

Principle Strain $e = \frac{\delta_l}{L}$

Strain Energy $= \frac{P\delta_l}{2}$

4.1.1 Sample Calculation for Left side race of 51326M Bearing

Area of Left Race = $\frac{\pi}{4}\left(d_1^{\,2} - d^{\,2}\right)$

$A = \frac{\pi}{4}\left(220^2 - 130^2\right)$

$A = 24.727 * 10^3 \text{mm}^2$

Principle Stress

$\sigma = \frac{P}{A} = \frac{230*10^3}{24.727*10^3} = 9.301 \text{N/mm}^2$

Axial Deformation or Displacement

$\delta_l = \frac{P*l}{A*E} = \frac{230*10^3*25}{24.727*10^3*2.08*10^5} = 1.1179*10^{-3}\text{mm}$

Principle Strain

$e = \frac{\delta_l}{L} = \frac{1.1179*10^{-3}}{25} = 4.4718*10^{-5}$

Strain Energy $= \frac{P\delta_l}{2A} = \frac{230*10^3*1.1179*10^{-3}}{2*24.727*10^3} = 0.01039 \text{N/mm}$

4.1.2 Formulas and sample calculations for Ball Part [7]

Area= $\frac{\pi}{4}$ (Diameter of Contact Surface with Race Groove) ^{2}

Diameter of Balls is 35mm and 50 mm and groove are 11 mm and 12 mm respectively for old and new bearings

Principle Stress $\sigma = \frac{P}{A}$

Deformation $\delta = \frac{P\ r^3}{E\ I}$

Where $I = \frac{\pi d^4}{64}$

Principle Strain $e = \frac{\delta}{r}$

Using these equations values of stress, strain, deformation and strain energy are calculated and plotted into the Results Tables

4.1.3 Sample Calculation for Ball part of 51326M Bearing

Area = $\frac{\pi}{4}$ (Diameter of Contact Surface with Race Groove) ^{2}

Area = $\frac{\pi}{4}(11)^2$

A = 94.985 mm^2

Principle Stress

$$\sigma = \frac{P}{A} = \frac{16428.57}{94.985} = 172.959 \text{ N/mm}^2$$

$$I = \frac{\pi d^4}{64} = \frac{\pi * (35)^4}{64} = 73.62 \text{ mm}^4$$

Deformation

$$\delta = \frac{P\ r^3}{EI} = \frac{16428.57*(17.5)^3}{2.08*10^5*73.62*10^3} = 0.0057496 \text{ mm}$$

Principle Strain

$$e = \frac{\delta}{r} = \frac{0.0057496}{17.5} = 3.28548*10^{-4}$$

$$\text{Strain Energy} = \frac{P\delta_1}{2A} = \frac{16428.57*0.0057496}{2*94.985} = 0.0284 \text{ N/mm}$$

4.2 Results

Table 3 Comparison of results obtained by Ansys and Pro-E for old and new bearing

Left Side Race

Property	Analytical		Software Analysis			
	Old Bearing 51326M	New Bearing 51426M	Ansys		Pro E	
			Old Bearing	New Bearing	Old Bearing	New Bearing
Maximum Principle Stress	9.301	5.4945	7.7204	4.3734	13.89	11.26
Displacement (X Axes)	1.1179E-03	9.7738E-04	0.0015071	0.0014808	0.0016399	0.0016318
Maximum Principle Strain	4.4718E-05	2.6415E-05	0.00004532	0.000026626	6.134E-05	0.00005011
Strain Energy	0.01039	2.6851E-03	0.57292	0.58714	0.0167	0.03325

Right Side Race

Property	Analytical		Software Analysis			
	Old Bearing 51326M	New Bearing 51426M	Ansys		Pro E	
			Old Bearing	New Bearing	Old Bearing	New Bearing
Maximum Principle Stress	8.9686	5.33259	5.4319	3.5906	11.50	5.282
Displacement (X Axes)	1.07795E-03	9.4858E-04	0.0014456	0.0012903	0.0013962	0.0012552
Maximum Principle Strain	4.3118E-05	2.5637E-05	0.000038212	1.1491E-07	0.00006191	6.911E-07
Strain Energy	4.833856E-03	2.52919E-03	0.77479	0.73595	0.001648	0.0004533

Ball

Property	Analytical		Software Analysis			
	Old Bearing 51326M	New Bearing 51426M	Ansys		Pro E	
			Old Bearing	New Bearing	Old Bearing	New Bearing
Maximum Principle Stress	172.959	203.46	171.18	128.34	91.22	1.066E2
Displacement (X Axes)	5.7496E-3	5.634E-3	0.0056303	0.0057205	5.6152E-3	5.7257E-3
Maximum Principle Strain	3.28548E-4	7.076E-4	0.0011577	0.0007779	1.802E-3	1.442E-3
Strain Energy	0.0284	1.7996	0.26988	0.6978	6.207	3.966

3.0 Conclusions

- In this dissertation work we designed some of the components of oil expeller such gear ratio, main shaft rpm, main shaft diameter and crushing chamber dimensions. Design of all these components matches with the components actually used in oil expeller machine. Therefore we can say that the design components used is safe.
- We also calculated the thrust force acting upon the thrust ball bearing, which is quite high which a bearing can withstand. Calculated life of existing bearing is about 5.2 months. Hence new bearing must be selected and hence we select bearing no 51426M which gives satisfactory life of 2.5 years by calculation. This bearing has been also replaced by the company and it is giving satisfactory result and has not failed yet.
- Further we have done static analysis of the Part of bearing on Pro-E wild Fire 4 and Ansys Workbench 11. In analysis results the Principle stress, Principle strain and axial deformation is reduced. Further the results obtained by different software are approximately matching the results obtained by the analytical solutions. From the comparison of the result data we found that results obtained by the Ansys Workbench are closer to the analytical results.

.

4.0 References

[1] Bamgboye and Adejumo, “Development of a Sunflower Oil Expeller”, Agricultural Engineering International: the CIGR E-journal, Manuscript EE 06 015, Vol IX. September, 2007

[2] Alessandra Maria Sabelli, "Design of a Press for Oil Extraction from Moringa Seeds for Haiti" Thesis, Department of Mechanical Engineering, Massachusetts institute of technology, January 2004

[3] D. Scott, “Bearing failure Diagnosis and Investigation”, Nationall Engineering Laboratory, East Kilbride (Gt. Britain), Vol Wear 25, Page No 199-213, 1973

[4] R.S. Khurmi and J.K. Gupta, “A Text Book of Machine Design”, Chapter 14 Shaft and Chapter 27 Rolling Contact Bearings, 14th edition, 2005

[5] ZKL Bearings, “Thrust Ball Bearings”, Manual Page No 174 183

[6] Flender Extruder Ltd, “Selection Process Manual”, K21-1.0 EN, 2005

[7] Indrajeet M. Jain, “Advanced Strength of Materials”, Tech Max Publication, Chapter 1, Page No 1-48

Hydrogen – An Alternative Fuel for 21st Century

Shaival R Parikh, Lecturer,
Mechanical Engg. Department,
Gandhinagar Institute of Technology.

***Abstract*—World is facing problems of Energy Crises and thus all countries are in search of alternative fuels. Alternative fuels in the sense that they can fully replace the existing non-renewable sources. Hydrogen is such fuel which can replace the conventional fuels. It can be produced from fossil raw materials or using processes like electrolysis of water, gasification of biomass etc. and thus is renewable energy source. It can be stored in gaseous, cryogenic liquid or metallic hydride form. Hydrogen has found wide use in domestic, industrial and transportation fields. It can be used in I.C. engines using Manifold Induction, Direct Introduction or as supplementary fuel for S.I. engines or in blending with diesel for C.I. engines. However, certain modifications in injector are required to prevent pre-ignition and back flash. As a fuel for I.C. engines it has high power output compared to conventional I.C. Engines. Use of Hydrogen permits negligible emissions and thus it is 'Environment Friendly'. However, Hydrogen can not be described as totally safe but the risk of accident can be reduced significantly by appropriate storage, handling and transportation methods. This paper describes the key features of Hydrogen which are beneficial for its use as an Alternative fuel for the coming generations.**

***Index Terms*— Electrolyser, Biophotolysis, Gasification, fuel cell technology.**

I. INTRODUCTION

Energy is the primary and most universal measure of all kinds of works by human beings and nature. Energy is an important input in all sectors of any country's economy. There are mainly two reasons responsible for energy crisis – high rate of increase in population of the world and increase in standard of living of peoples. Our efforts to fulfill high energy demands are making hazards to environmental factors, leading to major problems such as pollution, green house effects etc. On the other hand conventional energy sources are not able to supply till next 50 years.

So scientists and engineers are looking for alternatives which will fulfill our demands, environmental aspects and can sustain for long time. Today there are various renewable energy sources such as solar energy, wind energy, tidal energy under development. These sources are unlikely to contribute supply of energy. Hydrogen is also a renewable fuel as it can be produced again & again using various methods. Hydrogen technology is such a technology which can replace petroleum efficiently. It is a renewable source of energy and also its use has no major side-effects on environmental factors. It can be used in Fuel Cells or in I.C. engines directly with some modifications. Use of Hydrogen in I.C. engines permits negligible emissions which are not harmful to environment. If we want to implement use of any known renewable source in transportation system we have to use special equipments but Hydrogen can be used in conventional I.C. engines.

II. PRODUCTION OF HYDROGEN

Hydrogen can be produced using Fossil raw materials or using renewable energy.

1. By using Fossil Raw Materials

It involves heating of Hydrocarbons, steam and/or Oxygen which are then combined in a reactor.

- **Steam reforming of natural gas:** Natural gas consists of mainly methane with some heavier HC & CO_2. It is cheapest method of H_2 production involving following reactions.

$$CH_4 + H_2O \xrightarrow{\text{Catalyst, 700 – 1000 0C, 3- 25 bar}} CO + 3H_2$$
$$CO + H_2O \longrightarrow CO_2 + H_2 \uparrow$$

- **Thermal Dissociation:** It involves heating of hydrocarbons in absence of O_2 at high temp.

$$CH_4 \xrightarrow{\text{Pyrolysis}} C + 2H_2 \uparrow$$

- **Carbon Black & Hydrogen:** It uses carbon black (super pure carbon-soot). In a high temp. Reactor heat is supplied to split the hydrogen compound by plasma burner which utilizes recycled H_2 as plasma gas.
- **Plasmatron:** The plasma reformer operates in pyrolitic mode (thermal degrading of oxygen or air), turning carbon into soot, eliminating CO_2 formation. Its advantage is that it can use all forms of HCs. Main disadvantage of plasma reformer is its dependency on electric power.

2. By using Renewable Energy

- **Electrolyser:** It involves splitting of H_2O. Water is subjected to electrical power and the result is H_2 & O_2.

$$2H_2O + \text{Energy} \longrightarrow 2H_2 + O_2$$

The different types of electrolysers are Alkaline, polymer electrolyte membrane,steam electrolysers etc.

- **Photoelectrolysis:** Instead of first converting sunlight into electricity & then using electrolyser to produce H_2 & O_2 from water, it is possible to split H_2 and O_2 on the surface of photovoltaic cell. It eliminates the cost of electrolysers & increases system efficiency.

- **Biophotolysis :** It involves 'Photosynthesis' in which the first step is splitting of H_2 and O_2. In these blue green algae, *Anabaena* is used which can produce hydrogen, with an efficiency of 25%. These are single celled green algae able

to make the enzyme Hydrogenase & Ferredoxin. The exposure of these algae to sunlight and water yields a mixture of O_2 and H_2 gases.

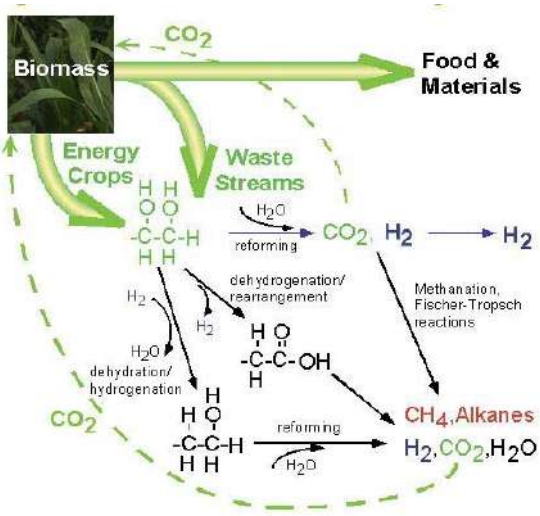


Figure. Hydrogen production from Renewable energy

III. Storage of Hydrogen

Hydrogen contains a lot of energy per unit weight but less energy per unit volume. This possess potential problem in terms of storing large amount of H_2 in small space. The traditional means of storage such as Cryogenic tanks and pressure tanks have improved dramatically and also numbers of new technologies of storage are under development.

1. **Gaseous storage:** At room temperature hydrogen is gaseous that can be stored in compressed gas cylinders similar to natural gas. Storage in gas cylinder under 200 bar pressure is feasible for stationary engines but it is very expensive and bulky as well as large amount of steels are needed to contain a small amount of hydrogen for locomotives. Compressed gas cylinders have pressure level of 20 MPa. A high pressure cylinder made of plastic composite structural materials with steel or aluminum liners are under development.
2. **Liquid storage:** Hydrogen can be turned to a Liquid by pressuring it at low temp. In a liquid Hydrogen tank at -253 ^{0}C, 845 cc of H_2 can be stored in 1 cc tank. Hence liquid hydrogen is preferred to compressed gas storage. This increases the cost as compared to gaseous storage. The energy content of liquid hydrogen is 2.75% greater than that of other hydrocarbon fuels. For storage of liquid hydrogen different storage tanks are developed consisting of metal linings rapped with fibers or aluminum foil layers separated by plastic foil. Therefore, liquid hydrogen is suitable for rockets, aircrafts and road transportation applications. The fig. shows a cryogenic liquid Hydrogen tank.
3. **Metallic Hydrides Storage:** Storage of hydrogen in metal hydrides is a promising line of development because of its smaller molecular size. The Hydrides of Lanthanum-Nickel ($LaNi_5$), Iron-Titanium (FeTi) and Magnesium-Nickel (Mg_2NiH_4) are used. For metals like Titanium the penetration is more. These Hydrides contain somewhat more hydrogen than an equal volume, but much less than an equal weight of liquid.

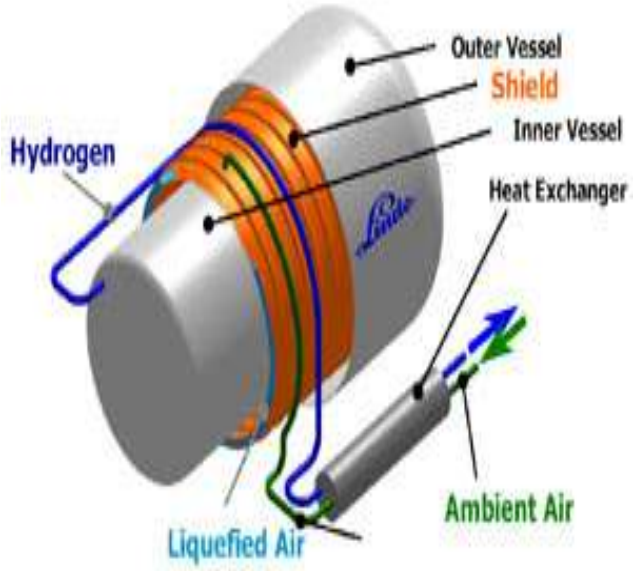


Figure. Hydrogen storage device

IV. Comparison of Hydrogen with other Conventional Fuels

Hydrogen has certain properties which makes it more useful compared to other fuels. The Chart given below explains the properties of hydrogen that makes hydrogen more advantageous than other fuels. The Heating value of Hydrogen ranges from 142 to 120 MJ/Kg which is much higher than other fuels. Also A/F Ratio for use of Hydrogen is 34.3 thus having more efficiency.

Fuel	Molecular Weight	Specific Heat KJ/Kg	Higher heating value MJ/Kg	Lower heating value MJ/Kg	(A/F) Ratio
Gasoline	110	1.7	43.3	44	14.6
Natural Gas	18	2	50	45	14.5
Propane	44.10	1.6	50.4	46.4	15.13
Toluene	92.14	1.1	42.5	40.6	13.50
Methanol	32.04	1.72	22.7	20	6.47
Ethanol	46.06	1.93	29.7	26.9	9.00
Hydrogen	2.015	1.44	142	120	34.3

V. Applications of Hydrogen

For Residential use: - To produce electricity for electric equipments and other applications in houses. It can be also

used for domestic cooking. The burners of domestic appliances should be modified if hydrogen is to replace natural gas. Direct combustion of hydrogen produces water only. Nitrogen oxides are formed in very small amount. Hydrogen will be useful in radiant space heaters because of possibility of flameless combustion on a catalytic surface.

For Industrial use: - Hydrogen is used in many processes either as fuel or chemical reducing agent. Hydrogen is also used instead of coal or coal derived gases to reduce oxide ores to the metals. The artificial fertilizers and petroleum industries are heaviest users of hydrogen. Sale of hydrogen has increased by 6% annually in the last five years because of the varied industrial applications.

For Transportation:- In the field of Transportation, Hydrogen can be used in following two ways - a) Fuel Cell Technology b) Hydrogen used in IC engines.

- **Application of Hydrogen in I.C. Engines**

Hydrogen can be used in both S.I. & C.I. Engines as a fuel. Hydrogen has extremely wide ignition limits. This allows a S.I. engine to operate on Hydrogen with very little throttling. Stoichiometric Hydrogen-air mixture burns 7 times faster than gasoline air mixture. Thus having higher engine speed and higher thermal efficiency. H_2 has high self ignition temp. but requires very little energy to ignite it. Hence it is highly prone to pre-ignition and back flash in S.I. engines. Adiabatic flame temperatures for Hydrogen is little lower than for gasoline. But rapid combustion allows very little heat loss to the surroundings and hence high, instantaneous, local temperatures are produced. This leads to high Nitric acid formation.

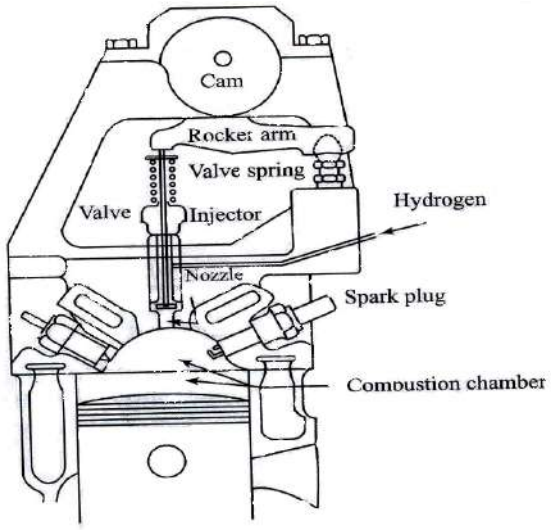


Figure. Hydrogen supply in I.C. Engines

VI. Advantages and Disadvantages of Hydrogen

Advantages

- As an energy carrier, hydrogen can be produced safely and abundantly from diverse renewable resources such as hydroelectricity, solar and wind power.
- It is environment friendly since no toxins, no SO_X, no NO_X fulfilling the Clean Air Act. Also Reduces greenhouse gas emissions, retarding 'Global Warming'.
- Hydrogen can be stored safely in special storage equipments, without harm to environment.
- Efficiency is improved at least 10 percent.
- Stronger national energy security, since dependency on foreign energy sources is removed.
- Leakage of Hydrogen fuel is not harmful to environment & human health.

` **Disadvantages**

- Cost: The major barrier regarding the use of Hydrogen is cost, which is quiet high. The high cost is due to different equipments. The expensive element is storage tanks in H_2 engines.
- Fuel Storage: Hydrogen contains a lot of energy per unit of weight while content of energy per unit volume is low. It possesses a potential problem in small space. Hydrogen can be stored in liquid, gaseous or in combined form. The storage of hydrogen is a topic of expensive research and development.
- Infrastructure: To launch the Hydrogen technology everywhere, sustainable development & better infrastructure must be achieved.
- Difficult to Refuel.

VII. Conclusion

In 21st century the world is facing problems like energy crisis, pollution, climatic changes etc. One of the best solutions for these problems is Hydrogen technology which has proved itself as promising candidate in the race of alternative fuels. It is efficient, environmental friendly, reliable & safe. The Hydrogen will surely be a 'Future Fuel' provided its technical challenges are overcome satisfactorily and economically.

References

[1] Nitin Karwa, Hydrogen as fuel for I.C. engines, Proceedings of the national seminar on "MODERN TRENDZ IN AUTOMOBILES."
[2] Seminar On Hydrogen Fuel Cell Engines and Related Technologies from College Of Desert
[3] Seminar on "Pathways To Clear Energy" By ChevronTexaco Technology Ventures LLC
[4] Renewable energy sources by G.D. Rai.
[5] I.C. Engines by John B. Heywood
[6] I.C. Engines by Mathur-Sharma.
[7] I.C. Engines by V. Ganesan.

Analysis of water flow inside the compartment of continuous Electro-deionization (CEDI) unit using Computational fluid dynamics

Krunal Patel[a], Girish R. Desale[b*], S.K. Thampy[b], Nitin G. Borole[b], Vikas J. Lakhera[a]

a:M.Tech Thermal Engg Student, Institute of Technology, Nirma University, Ahemdabad-382 481, INDIA
E-mail: krunal_7455@yahoo.com
b: Desig*n Drawing and Workshop Engineering Department, Central Salt and Marine Chemicals Research Institute, Council of Scientific and Industrial Research (CSIR), G.B. Marg, Bhavnagar-364 002,Gujarat, India.*

Abstract

The method of producing ultra water is a hybrid process combining of three technologies- Reverse Osmosis (RO), Continuous Electro deionization (CEDI) and Deionization. The RO product water (25-40 µS/cm) is supplied to CEDI unit in two inlet streams; one for treated compartment (product water) and another for concentrated compartments (rejected water). Finally the treated water of CEDI unit is passed through the polishing unit (Deionization process) and ultra pure water of 18.2 MΩ.cm is collected for further use. The ultra pure water has number of application in industries as well as in academic institutes and research laboratories. Looking to the interest of multinational companies towards Research and Development, most of the industries in India requires ultra pure water system for research work. Thus from the market demand, it appears that ultra pure water system has lot of scope in the Indian market as well as in abroad. The most important part of the ultra pure water system is CEDI unit, in which the R.O product water is passed through treated and rejected compartment (separated by ion exchange membrane). The ions will get transfer from treated compartment to rejected one due to applied voltage enhance the transfer of ions the compartments are filled with ion exchange resin. Thus CFD may be the most effective tool to analyze the flow behavior and its pattern inside the compartment.

1. INTRODUCTION

The demand for high-purity industrial process water is rapidly increasing. Large quantities of water of different quality require-ments are used as boiler feed water, in the production of semiconductors, and in chemical and biochemical laboratories. Conventional systems for high-purity water production contain several process steps in series such as microfiltration and reverse osmosis. As a final demineralization step ion-exchange with mixed beds is applied where the ion-exchange resin has to be regenerated or replaced in certain time intervals.

. Thus, application of CEDI as a final demineralization step after reverse osmosis desalination is currently a cost-effective alternative to the previously used mixed-bed ion-exchange in the production of high-purity process water. A number of manufacturers supply a rapidly growing market with water

Today, ion-exchange is more often replaced by a process referred to as continuous electro deionization (CEDI). Compared with ion-exchange the CEDI process provides a product water of equal or better quality and is less costly since no separate regen-eration step or replacement of ion-exchange resin is required demineralization systems containing CEDI, where the CEDI stage varies significantly in its basic concept and in stack design.

The first commercially available CEDI modules [3] were similar to a conventional electro dialysis stack with cat ion-and anion-exchange membranes in alternating series, forming

individual dilute and concentrate compartments between two electrodes. In contrast to conventional electro dialysis, the dilute-

ate compartment in CEDI stack was filled with a mixed-bed ion-exchange resin.

Several modifications of the original CEDI stack design have meanwhile significantly improved the efficiency of the process. The electrical resistance of the stack was substantially reduced by filling also the concentrate compartment with a mixed-bed ion-exchange resin.

The changes applied have lowered the overall process and maintenance costs, the energy consumption and the susceptibility to scaling and fouling, and increased water recovery as well as the ability to remove weakly dissociated contaminants. The main concepts of the CEDI processes are described in detail in the corresponding literature and shall be briefly reviewed in the following section.

Present Study

Today, ion-exchange is more often replaced by the process referred to as continuous electro deionization (CEDI). Compared with ion-exchange the CEDI process provides a product water of equal or better quality and is less costly, since no separate regeneration step or replacement of ion-exchange resin is required. Thus, application of CEDI as a final demineralization step after reverse osmosis desalination is currently a cost-effective alternative to the previously used mixed-bed ion-exchange in the production of high-purity water. A number of manufacturers are interested in the rapidly growing market of water demineralization systems containing CEDI, where the CEDI stage varies significantly in its basic concept and in stack design are used. Hence, CEDI unit having capacity of 10 lit/hr developed at CSMCRI Bhavnagar is considered for CFD analysis. The detailed specifications and drawings of the CEDI were obtained from Central Salt Marine Research Institute (CSMCRI), Bhavnagar at the request of INFUSIL INDIA PVT. LTD,Bangalore. The configuration of CEDI consists of 5 numbers of treated compartment; the salient features of the CEDI compartment are given in Table. 2.1. The details of CEDI treated compartment only are shown in Fig. 2.4.

2 COMPUTATIONAL MODELING

The computational modeling, meshing of the geometry & specifying zones are initial steps in CFD analysis. These three steps are performed in GAMBIT, which is a pre-processor of FLUENT. The complete methodology applied for the computational modeling, meshing & CEDI compartments are described in following paragraphs.

2.1 GAMBIT

2.1.1 Compartment Geometry Creation

The geometries of the different components (i.e. inlet section, outlet section and active area of CEDI comportment) are created in the pre-processor called GAMBIT. The hierarchy of geometric objects in GAMBIT is Vertices → Edges→ Faces. This approach is called Bottom-up approach.

Initially, the geometry of the component i.e. CEDI compartment is created and its details are as follow:

CEDI compartment module plate

The CEDI compartment is divided into three different sections and three sections are of different sizes and the steps are as follow:

a) Create three different sections.

i. Three inlet nozzles
ii. The middle section of compartment/plate having channel (Active area)
iii. Three outlet nozzles (the dimensions of all nozzles at inlet and outlet are same)

b) Create the number of faces between successive sections

All the vertexes of inlet section, channel section (active area) and outlet section are joined together. The complete CEDI compartment design after joining all the edges converted into faces in GAMBIT is shown in figure 2.1.

Domain	Size , mm	Type of grid
Inlet nozzle (Trapezoidal)	16 x 6 x 8	Tri elements with pave scheme. Interval count: 10 growth rate: 1 size limit: 58
Outlet nozzle (Trapezoidal)	16 x 6 x 8	Tri elements with pave scheme. Interval count: 10 growth rate: 1 size limit : 58
Compartment channel	250 x 65	with size function having growth rate: 1 size limit: 250

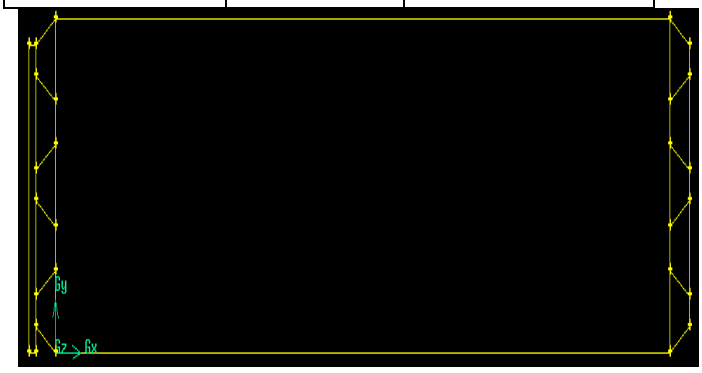


Figure 2.1 CEDI compartment in GAMBIT

2.1.2 Grid Generation

One of the most important and time-consuming challenge in the CFD simulation process is the generation of the computational grid or mesh. As the geometry is complex the unstructured grid is used with square and tetrahedral elements. The size of the element during meshing defines the ability to solve the problem and its accuracy. The parameter size function controls the size of mesh intervals for edges and meshing elements at faces. Size functions are similar to boundary layers in that they control the mesh characteristics in the proximity of the entities to which they are attached. They differ from boundary layers with respect to the manner in which they are defined and the manner in which they control the mesh. Whereas boundary layers prescribe specific mesh patterns and the sizes of mesh elements within those patterns.

The steps for grid generation:

i. Mesh the inlet faces using tri elements with pave scheme.

ii. Mesh the channel faces using quadrilateral elements with map scheme.

iii. Mesh the outlet faces using tri elements with pave scheme.

iv. Examine the mesh quality for equi-size and equi-angle skeewness elements.

The details of type of the grids, it is size and application to the respective domain of the compartment in GAMBIT is given in Table.2.1. The geometry generated using this different goods in GAMBIT is shown in Figure 2.2.

Table 2.1 Details of types of grids

Number of nodes: 153777 Number of mesh elements (cell) 152611

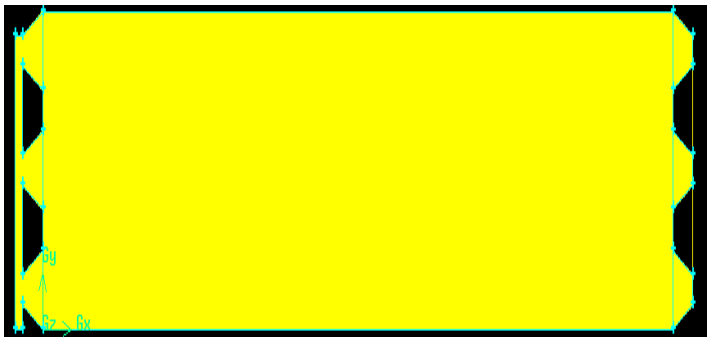


Fig. 2.2 Grid for inlet, active area & outlet of the compartment.

2.1.3 Zone Specification

Zone specifications define the physical and operational characteristics of the model at its boundaries and within specific regions of its domain. The zone specifications are classified as Boundary types and Continuum types. Before exporting the mesh geometry to FLUENT, it is required to specify the zone types in GAMBIT and described in detail as;

i. Boundary type

Boundary-type specifications, such as VELOCITY INLET and WALL define the characteristics of the model at its internal and external boundaries respectively. The boundary types used in the present investigation are specified as;

a. At inlet of plate: Velocity Inlet
b. Compartment Walls : no slip
c. Porous Jump in (50% slit opening)
d. Porous Jump out (50% slit opening)
e. Other than the water flow edges considered as default interior
f. At the plate outlet: Static Pressure Outlet

ii. Continuum type

Continuum-type specifications, such as FLUID and RESIN define the characteristics of the model within specified regions of its domain. The continuum types used in the present investigation are specified as;

a. Inlet nozzle : Fluid
b. Active area : Porous media fluid (resin),
c. Outlet nozzle : Fluid

After completing the modeling, meshing and specifying zones with boundary conditions in GAMBIT, the entire mesh is exported to 'FLUENT 5/6 Solver' for simulation.

2.2. SIMULATION IN FLUENT

The predictions of the complex flow inside the compartment of CEDI unit can be achieved significantly using CFD. These predictions are supportive while designing the CEDI unit compartment. This article describes the 2-dimensional (2d) simulation of flow in the compartment. A commercial available 2d Navier-Stokes code called 'FLUENT' with laminar model is used to simulate the flow conditions. Finite Volume method (FVM) is used for the discretization of governing equations. To avoid the more complexity, better understanding of the flow profile and to have better predictions, certain assumptions made during the simulation of flow conditions are given below as;

2.2.1 Assumptions

i. Flow is steady state.
ii. Fluid is incompressible.
iii. Fluid properties are constant.
iv. No vapor is present in the water, i.e. single phase flow.
v. There is no leakage in the plate.
vi. The surface of all the components is hydraulically smooth.
vii. No mass (fluid and/or ions) transfer from the compartment through the active area (i.e. membrane).
viii. Resin bed is uniformly distributed in the compartment.
ix. Equal pressure drop in each treated compartment.

2.3 Solution Technique

Generally, single precision version or double precision version computational facilities available with FLUENT's are used for simulation. In the present investigation, the double precision version computational facility is used due to more accuracy and capacity to round-off errors. The different steps involved in computation are given below as;

i. The grid is read and checked in FLUENT.
ii. The segregated implicit solver with absolute velocity formulation is used for the computation. The cell-based gradient scheme is used which is more accurate than the cell-based scheme for structured meshes, most notably for quadrilateral meshes.
iii. Water with following properties is used as a working fluid.

    a. Density = 998.2 kg/m^3

b. Viscosity = 0.001003 kg/m-s

iv. Porous media with following properties.

a. Density 700 kg/m^3

b. Viscosity 1 kg/m-s

v. Porous Jump boundary condition applying at outlet of inlet nozzle and inlet of outlet nozzle.

a. Face permeability (m^2) : $1x10^{10}$

b. Porous medium thickness (m) : 0.001

c. Pressure-jump Coefficient (C_2)(1/m) : 3125 (at 50% opening)

vi. Following convergence criteria is used for different equations.

a. Continuity equation 0.0001

b. Momentum equations 0.0001

vii. A SIMPLE scheme is used for the pressure velocity coupling.

viii. For pressure correction, STANDARD scheme is used which is suitable for static pressure problems.

ix. For momentum and turbulence, first order upwind scheme is used.

x. Initially, following under relaxation factors are used.

a. Pressure 0.5

b. Momentum 0.2

Now the boundary conditions are required to define and apply to the problem through FLUENT. These boundary conditions are defined either by using experimental data or from calculations.

2.4 Experimental Set up and Results

The photographic view of the experimental set up consists of ultrapure water unit equipped with manometer is shown in figure 2.3. A conventional Bernoulli equation was used to calculate the pressure drop across the channel [Darcovich et al., 1996]. Similarly, U-tube manometer is used to measure the pressure drop of the entire CEDI unit. It is observed that the total pressure drop in the CEDI unit is 27.575 KPa the same was also measured with the pressure dial gauge and was around 4 psi. The water flow rates form the treated compartments at inlet and outlet is measured using measuring flask and stop watch. It is observed around 10 liter/hr from the CEDI unit.

2.3 Practical set-up

Practically, it is very difficult to calculate the pressure drop of each compartment, hence equal pressure drop in each compartment is assumed. Thus the pressure drop (ΔP) for individual compartment is calculated dividing the total pressure drop by number of compartments.

Thus,

Total Pressure drop = 27.575 KPa

Number of compartments = 5

Pressure drop in each compartment = 27.575 /5 = 5.515 KPa.

Similarly, from the measured flow rate and the detail dimesions of plate, velocity at different sections is calculated using equation 2.1.

Outlet Velocity v_2 (i.e. from the tube of inner diameter 6mm)

$$\dot{m} = \rho v A \qquad (2.1)$$

Where m is mass flow rate 0.003 Kg/Sec, ρ is density of water 1000 Kg/m^3, v is velocity in m/s, area of tube is m^2.

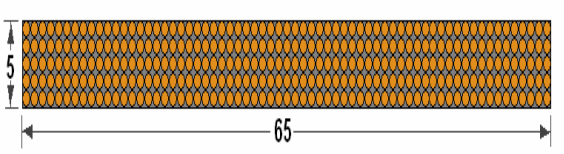
$$A = \frac{\pi}{4} d^2$$
$$A = \frac{\pi}{4} 6^2$$
$$A = 28.27 mm^2 = 0.02827 m^2$$

Thus, v_2 =0.09825m/s

Similarly, velocities and pressure at inlet and outlet for different compartments are calculated and given in table 4.3

2.5 Boundary conditions

The compartments of the CEDI unit are packed with ion exchange mixed bed resin. The average diameter of the resin ball is 1mm. The mixed bed resin is filled inside the compartment however, as discussed in section 2.2.1 it is assumed that the resin balls are uniformly distributed in the compartment as shown in figure 2.4. Thus the compartment is considered as porous media with porosity (n) for packing angle (θ) 90°.



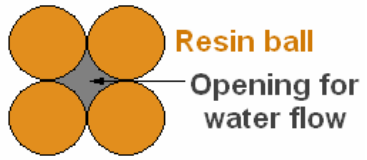


Figure2.4:Arrangement of resin balls inside the compartment

The flow inside the compartment is through porous media and hence Idelchik Reynolds Number equation 2.2 is used [LMNO engineering home page]. To calculate the Idelchik Reynolds Number, the porosity of the media is calculated using equation 2.3. Also from figure 2.4 the opening area for passage of water is calculated from simple equations and total area of opening for water flow is calculated. From the mass flow rate of water at 10 liter/hr, velocity of water inside the porous media is calculated.

Total opening area (A) = $2.146 \times 10^{-7} \times 5 \times 65$

$= 6.97 \times 10^{-5} m^2$

Using equation 2.1 the velocity of water inside the compartment (active area)is calculated as;

$$\frac{10}{3600} = 1000 \times v \times 6.97 \times 10^{-5}$$
$$v = 0.03985 m/s \qquad (2.1)$$

Idelchik Reynold Number for porosity, $\mathrm{Re} = \frac{0.45 v d \rho}{\mu (1-n)\sqrt{n}}$

(2.2)

Where, ρ is density of water 1000 Kg/m^3, v is velocity in m/s, d is resin diameter in m, μ is dynamic viscosity of water 0.848×10^{-3}N-s/m^2, n is porosity and calculated using equation 4.3.

$$Porosity, n = 1 - \frac{\pi}{6(1 - cos\theta)\sqrt{1 + 2cos\theta}} \quad (2.3)$$

n = 0.476

$$Calculated\ Velocity = 0.03985\ m/s$$

Idelchik Reynolds Number = 58.71

Finally, the Idelchik Reynolds Number has been calculated using equation and its value is 58.71. Based on these calculations and Reynolds Number the flow inside the compartment is consider as laminar flow and laminar flow model is selected to compute the flow characteristics inside the compartment.

The boundary conditions are the most critical parameters which governs the solution, hence due care has been taken while specifying the boundary conditions. In present investigation following boundary conditions are used.

a. *Inlet Boundary conditions:*

Inlet boundary conditions, namely, velocity inlet, initial gauge pressure and operating pressure are calculated using appropriate equations are specified above experimental data. The inlet boundary conditions for all the compartments are given in Table 2.2.

b. *Outlet Boundary conditions:*

At the exit face of the outlet nozzle of respective compartment, static pressure outlet (outlet gauge pressure) is specified and given in Table 2.2

Table 2.2 inlet and outlet boundary conditions for different compartment

Compartment Number	**Inlet boundary conditions**			**Outlet boundary conditions**
	Initial gauge pressure	**Inlet Velocity, m/s**	**Operation pressure, KPa**	**Outlet gauge pressure, KPa**
1	0	0.2005	126.14	27.575
2	0	0.1799	120.63	22.25
3	0	0.15956	115.11	16.74
4	0	0.1391	109.59	11.22
5	0	0.1187	104.08	5.52

After modeling, meshing and specifying the zones for each compartment, entire geometry was exported into the FLUENT for solving the problem. The data files of the results are stored in the fluent and the image generated are presented & discussed in next chapter.

2.5 Result and Discussion

The product water from Reverse osmosis (R.O.) having TDS around 25 ppm is fed in two separate streams to treated and rejected compartments of CEDI unit. The quality of ultrapure water or the efficiency of the CEDI unit depends on the parameters namely, properties of membrane, applied voltage, thickness of compartment, active area, effective area of membranes, mass flow rate of fed water, velocity of water in the compartment, design of CEDI unit, R.O. product water quality, etc. Chemical engineers, and finally the ion exchange membrane developed at CSMCRI have been used in present investigation.

The flow inside the CEDI unit is complex and involves lot of fluid mechanics, the effect of parameters namely; pressure drop, water flow rate (distribution) and its velocity can be investigated using numerical methods. These parameters significantly affect the effective membrane area from the active area of compartment for ion transportation. The mass flow rate and velocity of water cannot predict the distribution of water

inside the porous compartment (i.e. filled with mixed bed resin). The water flow inside the compartment may follow the least resistance path without covering the maximum area inside the compartment. The narrow stream of water due to least resistance path reduces the contact time and membrane contact area inside the compartment. Due to this, ions will not have sufficient time to get transport from treated compartment to rejected one. Thus the improvement in the design of inlet and outlet section of each compartment is responsible for flow distribution vis-à-vis maximize the effective area of membrane which significantly improves transportation of ions.

3.1 CFD results on flow characteristics

The CFD solutions for the velocity and flow distribution (effective area of membrane) with Navier-Stroke equations were obtained for five different compartments of the CEDI unit. The velocity and flow contour along the length of the compartments filled with resin (porous material) generated using FLUENT and presented in Figure 3.1(a-e) for all five compartments respectively. Velocities at different zones of compartment are computed using the CFD and validated with theoretically calculated values. In the present approach, based on the velocity contour efforts have been made to identify the area of flow for maximum transportation of ions with respect to mass flow rate of water.

3.1.1 Flow velocity in active area

In order to investigate the effective area of membrane, the velocity of water and its distribution plays important role. Figures 3.1 (a-e) shows the velocity contours for all five compartments respectively. It is observed from figure 3.1(a-e) that the velocity contours are different for first compartment as shown in Figure 3.1 (a) compare to the remaining four as shown in Figure 3.1 (b-e). The range of minimum and maximum velocity observed in first compartment is from 0.0174 m/s to 0.139 m/s. while the range of velocity for other compartments is from 0.02 m/s to 0.06 m/s. The water flow inlet for first plate is from single tube and further distribution in parallel to three trapezoidal inlet sections, where unequal distribution of water takes place. While in case of all other four compartments the inlet flow of water to the compartment is from the outlet (discharge) water from the three trapezoidal section of the previous compartment.

3.1.2 Flow pattern & Effective membrane area

The flow pattern inside the compartment and effective membrane area from the active area are the most important parameters responsible for the efficiency of CEDI unit. The uniform distribution of incoming water inside the compartment and its continuous in touch (contact) with the membrane area enhance the transportation of ions from treated compartment to concentrate. This can be achieved significantly by optimizing the velocity of water inside the compartment. Reducing the velocity of water below certain limits. i.e. flow rate of water may reduce the productivity of CEDI unit. The increase in the effective area of membrane and approaches ideally equal to active area of compartment increase the efficiency of the CEDI unit. However, effective area of membrane cannot be equal to the active area due to fluid mechanics problems associated with the flow of water inside the closed porous compartment. While increasing the velocity beyond certain limit may reduce the quality of ultrapure water as the ions may not get sufficient time to cross the membrane boundary.

It is postulated that the velocity contours for the compartments as shown in Figure 3.1 (a-e) also depicts the flow pattern inside the compartment. The compartments are filled with resin balls and hence water finds its own path to flow from the gaps between the resin balls. Due to these micro-channels get formed inside the compartment and the flow is considered as laminar as discussed in previous chapter. The average velocity in first compartment (see Figure 3.1 (a)) is around 0.017 m/s and uniform across the length and width of compartment. However, all other four compartments show different flow pattern compare to first compartment (see Figure 3.1 (a-e)).

Similar flow pattern has been observed for remaining other compartments (see Figure 3.1 (b-e)). In the active area two flow paths of different velocity contours have been observed. The flow path having minimum velocity around 0.02 m/s has been observed and initiated at the gap between the two trapezoidal sections at inlet side and ends

at the other side. The area covered by the velocity contours of 0.02 m/s velocity is measured around 21.6% of the active area. While three similar water flow paths of velocity around 0.06-0.08 m/s observed in the active area. The area covered by these flow path contours is around 78.4% of active area (see Figures 3.1 (b-e)). The width of these flow paths across the length of compartment is same and equal to the width of trapezoidal section outlet. Thus the 21.6% area of active area covered by water velocity contours of 0.02 m/s may not transfer the ions effectively, while the remaining area around 78% play important role in transportation of ions.

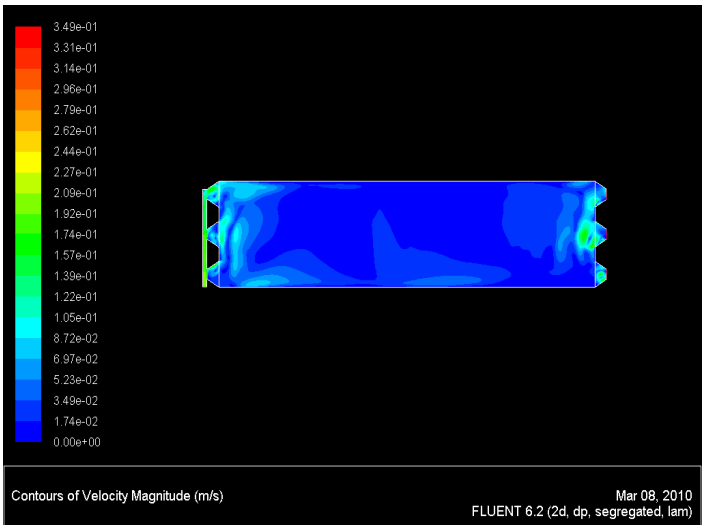


Figure 3.1 (a)

Figure 3.1 (b)

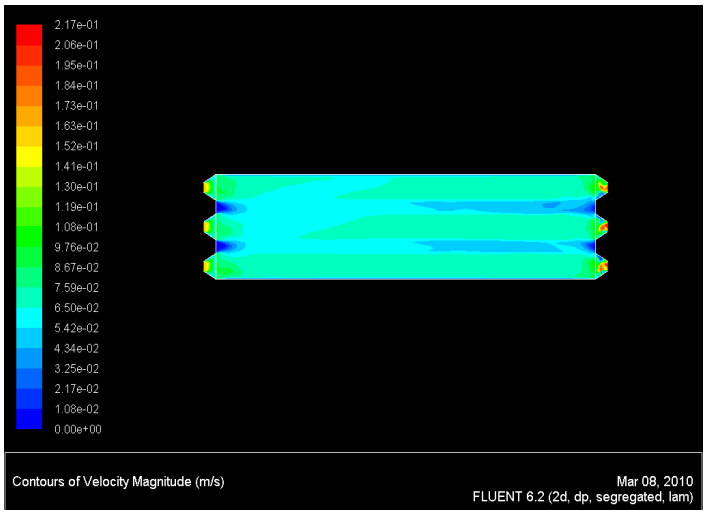
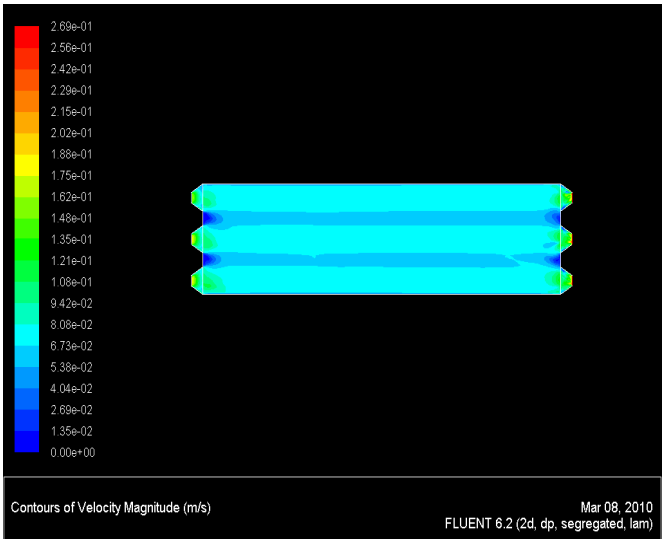


Figure 3.1 (c)



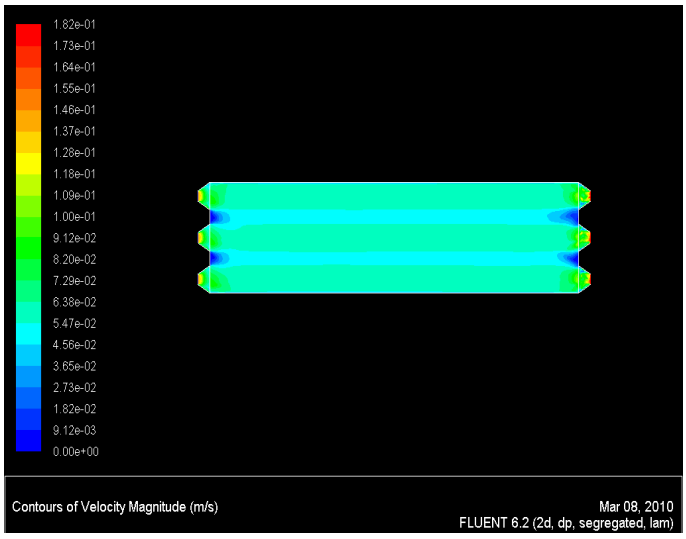


Figure 3.1 (d)

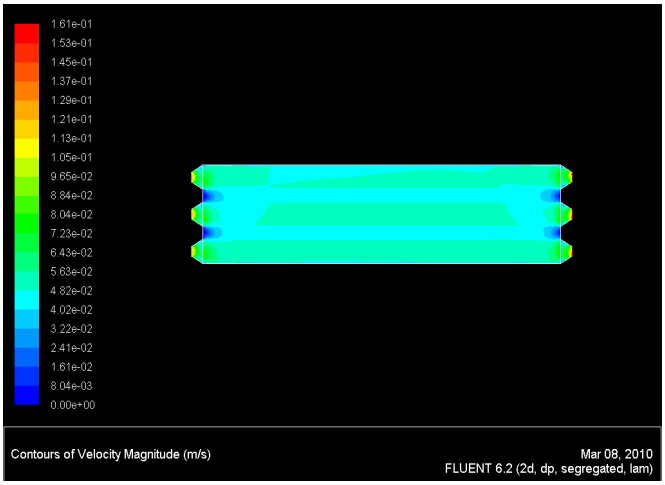


Figure 3.1 (e)

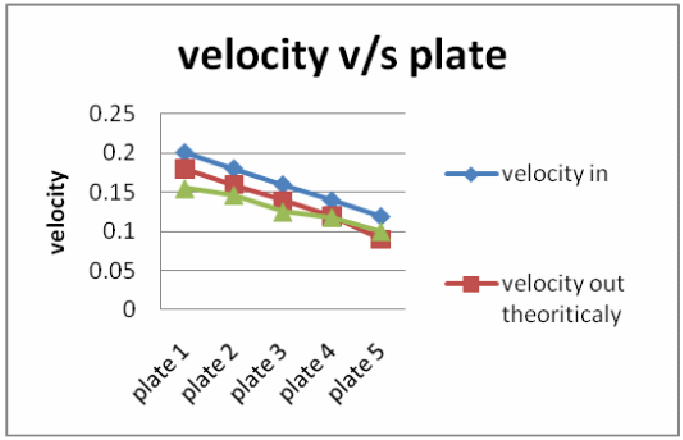


Figure 3.2

3.2 validation of result

As discused in chapter number four the inlet velocity for the respective compartment is calculated and given in Table 2.2. and presented grapicaly in Figure 3.2. The corrosponding outlet velocity are theoriticaly calculated and simillarly presented in Figure 3.2. The velocity contour for all compartments represents the velocity and flow pattern (path) as shown in Figure 3.1 (a-e) simillarly for corrosponding compartments the X_Y plot for velocity inlet and outlet where generated (not presented here) and average computed outlet velocity using CFD has been ploted in Figure 3.2. Variation between theoritical calculated and computed velocities has been observed in figure.

The maximum daviation for theoriticaly calculated and computed outlet velocity for first compartment is observed around 13%. This highest daviation may be due to parralel inlet flow condition for the first compartment. While for remanig four compartments the daviation in outlet valocity is observed arround 5%. Simillarly all compartment velocity observed considering porous media fill inside the compartment is computed between 0.06-0.08 m/s. The theoritical porous media velocity calculated is 0.03985 m/s.

Normally, the major water flow inside the compartment is along the length of compartment, anion and cation membranes are arranged alternatively to form the compartments. The ions are transferred from treated compartment to rejected (concentrate compartment) in the direction perpendicular to the plane of figure due to electrical energy potential applied to the electrodes. Along with the ions few quality of water also gets transferred through membranes. However, in the present simulation the membrane are assumed as walls in addition to this the resin filled inside the compartment is assumed uniformly distributed as shown in Figure 2.4.

4. Summary

The water flow characteristics of only treated compartments of CEDI unit (developed at CSMCRI Bhavnagar) have been investigated using CFD. The compartments are formed by alternate arrangement of cation and anion exchange membranes. The space between the membranes is filled with ion exchange mixed bed (MB-12) resin and consider as porous media in the present investigation. The R.O. product water is fed to the concentrate and treated compartments. The geometry of the inlet and outlet section of the compartment for water flow plays important role to decide the effective area of membrane and thus affects the performance of the ultrapure water unit.

The flow inside the porous compartment is considered as laminar based on the Idelchik Reynolds number calculation. Boundary conditions for the present investigations are decided based on the experimental results and assumptions. The

simulation techniques available with FLUENT based on Navier-stokes code is used to compute the flow characteristics of the compartments. Based on this analysis and theoretical calculations following conclusions can be drawn.

- The theoretically calculated flow velocity for the porous media is 0.04 m/s. The velocity contours generated using FLUENT shows uniform distribution of water inside the compartment with flow velocity in the range of 0.04-0.08 m/s, which is well in agreement with the theoretical calculations.
- Except first compartment outlet velocity, for remaining four compartments the theoretically calculated and computed values are within 5% error.
- The flow distribution of water inside the compartment decides the effective area of membrane from the active area which further decides the transportation of ions and quality of ultrapure water. The area with minimum flow velocity of water inside the compartment represents the dead zone or immobility area for ions while area with moderate flow velocity of water actively take part in ions transportation. From the velocity contours for the compartments with present inlet and outlet geometry for water flow shows around 78% area actively takes part in transportation of ions.
- Based on these conclusions it can be summarized that the present design for the compartment shows 78% utilization of membrane (i.e. effective area of membrane).

4.1 Future Scope of Work

The present problem requires good understanding of CEDI unit with its flow arrangement and fluid mechanics involved. Since the compartment if filled with mixed bed resin requires special application such as flow through porous media available with FLUENT. In the present investigation, flow characteristics for only treated compartment are analyzed using FLUENT. The following studies need to address the problems associated with CEDI unit with greater attention for detail and systematic approach require for its cost effectiveness.

- Two dimensional approaches is used in the present investigation, additional efforts are required to solve the problem with three dimensional mode.
- The flow characteristics for both treated and rejected compartments may be carried out with their common consequence on the CEDI unit.
- Based on the studies for different inlet and outlet geometries of the compartment the best design can be proposed for optimization of the CEDI unit.

Acknowledgements

The work was conducted by Central Salt Marine Research Institute, Bhavnagar.

BIBLIOGRAPHY

1. Acheson, D. 1990 Elementary Fluid Dynamics. Claredon Press, Oxford.
2. Belfort, G. 1972 an experimental study of electro dialysis hydrodynamics. Desalination 10 (221-262).
3. Andrej Grabowski', Guiqing Zhang, Heiner Strathmann, Gerhart Eigenberger, April 2007, Production of high-purity water by continuous electrodeionization with bipolar membranes: Influence of concentrate and protection compartment. www.sciencedirect.com
4. *TED PRATO, CHRISTOPHER JAY GALLAGHER* Using EDI to Meet the Needs of Pure Water Production, Ionics Incorporated, 65 Grove Street, Watertown, MA 02472, US,IWC-00-47 Report
5. Carl-Ola Danielsson[1], Anna Velin[2], Anders Dahlkild[1] & Marten Behm' Flow Distribution Study of a New Electro dialysis Module. www.mech.kth.se/thesis/2004/lic/lic_2004_**carl-ola**_danielsson

Palm Oil – An Alternative Fuel for I. C. Engines

Milan J Pandya,
Lecturer, *Mechanical Engg. Department,*
Gandhinagar Institute of Technology.

***Abstract*—Man in his lifetime, uses energy in one form or the other. In fact what ever happens in nature, results out of the conversion of energy in one form or the other? The extensive usage of crude oil has resulted in an energy crisis and there is a need to develop methods of optimal utilization, which will not only ease the crisis but also preserve the environment. Several other fuels such as hydrogen, nitrogen, nuclear, various ranges of vegetable oils are being tried now-a-days. The main objective of this paper is to explain the use of Palm oil as a fuel in Internal combustion engines and how the oil is produced from the palm fruit. It is observed that palm oil has higher density than petrol or diesel; hence it cannot be utilized directly as petrol or diesel. There are methods of using palm oil, straight vegetable oils or palm biodiesel.**

***Index Terms*— Palm kernel oil, threshing, clarification, kernel recovery, straight vegetable oils, transesterification, ELSBETT conversion technology.**

I. INTRODUCTION

Palm oil and palm kernel oil are edible plant oils derived from the fruits of palm trees. Palm oil is extracted from the pulp of the fruit, while palm kernel oil is derived from the kernel (seed) of the oil palm. Palm oil is naturally reddish because it contains a high amount of beta-carotene. Palm oil and palm kernel oil are three of the few highly saturated vegetable fats. Palm oil is semi-solid at room temperatures. Palm oil contains several saturated and unsaturated fats in the forms of glyceryl laurate (0.1% saturated), myristate (0.1% saturated), palmitate (44% saturated), stearate (5% saturated), oleate (39% monounsaturated), linoleate (10% polyunsaturated) and linolenate (0.3% polyunsaturated). Palm kernel oil is more highly saturated than palm oil.

II. PRODUCTION OF PALM OIL

Conversion of crude palm oil to refined oil involves the removal of the products of hydrolysis and oxidation, colour and flavor. After refining, the oil may be separated into liquid and solid phases by thermo-mechanical means. Extraction of oil from the palm kernels is generally separate from palm oil extraction and is often carried out in mills that process other oil seeds (such as groundnuts, rapeseed and cottonseed). The stages in this process comprise grinding the kernels into small particles, heating (cooking) and extracting the oil using an oilseed expeller or petroleum-derived solvent. The oil then requires clarification in a filter press or by sedimentation.

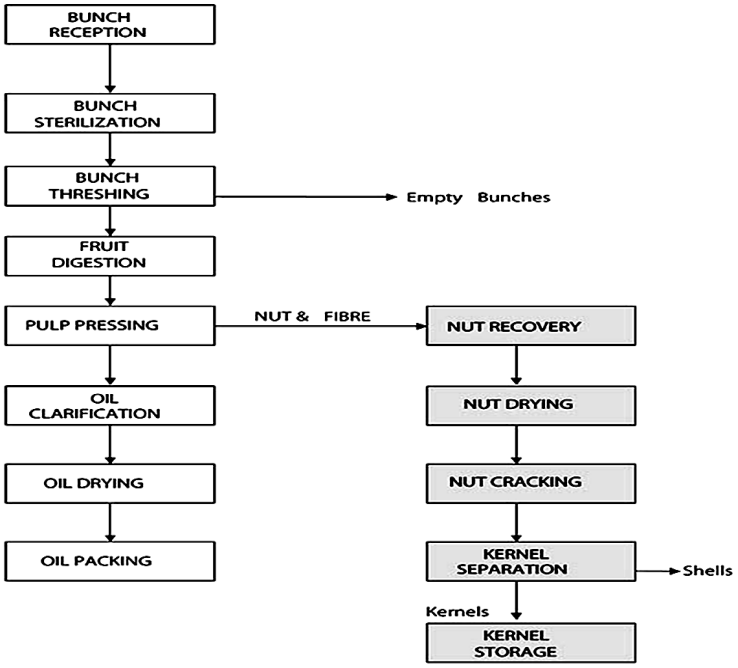


Figure. Flow Chart of Palm Oil Production.

1. **Bunch reception: -** Fresh fruits arrive from the field as bunches or loose fruit. The fresh fruit is normally emptied into wooden boxes suitable for weighing on a scale so that quantities of fruit arriving at the processing site may be checked. Large installations use weigh bridges to weigh the materials in trucks.

2. **Threshing (Removal of fruit from the bunches):-** The fresh fruit bunch consists of fruit embedded in spikelets growing on the main stem. Manual threshing is achieved by cutting the fruit-laden spikelets from the bunch stem with an axe or machete and then separating the fruit from the spikelets by hand. In the mechanized system, a rotating drum or fixed drum equipped with rotary beater bars detach the fruit from the bunch, leaving the spikelets on the stem.

3. **Sterilization of bunches: -** During this process, it is important to ensure evacuation of air from the sterilizer. Air not only acts as a barrier to heat transfer, but oil oxidation increases considerably at high temperatures, hence oxidation risks are high during sterilization. Over-sterilization can also lead to poor bleach ability of the resultant oil. Sterilization is also the chief factor

responsible for the discoloration of palm kernels. This leads to poor bleach ability of the extracted oil and reduction of the protein value of the press cake.

4. **Digestion of the fruit: -** The digestion is the process of releasing the palm oil in the fruit through the rupture or breaking down of the oil-bearing cells. The digester commonly used consists of a steam – heated cylindrical vessel fitted with a central rotating shaft carrying a number of stirring arms.

5. **Pressing (Extracting the Palm oil):-** There are two distinct methods of extracting oil from the digested material. One system uses mechanical presses and is called the dry method. The other is called the wet method uses hot water to remove the oil. In the dry method, the objective of the extraction stage is to squeeze the oil out of a mixture of oil, moisture, fiber and nuts by applying mechanical pressure on the digested mash.

6. **Clarification and Drying of oil: -** The main point of clarification is to separate the oil from its entrained impurities. The fluid coming out of the press is a mixture of palm oil, water, cell debris, fibrous material and non oily solids./ because of the non oily solids the mixture is very thick (viscous). Hot water is therefore added to the press output mixture to thin it. The dilution (addition of water) provides a barrier causing the heavy solids to fall to the bottom of the container while the lighter oil droplets flow through the watery mixture to the top when heat is applied to break the emulsion (oil suspended in water with the aid of gums and resins). The water is added in the ratio of 3:1.

7. **Oil storage: -** In large-scale mills the purified and dried oil is transferred to a tank for storage prior to dispatch from the mill. Since the rate of oxidation of the oil increases with the temperature of storage, the oil is normally maintained around 50 deg C, using hot water or low pressure steam heating coils, to prevent solidification and fractionation. Iron contamination from the storage tank may occur if the tank is not lined with a suitable protective coating.

8. **Kernel recovery: -** The residue from the press consists of a mixture of fibre and palm nuts. The nuts are separated from the fibre by hand in the small-scale operations. The sorted fibre is covered and allowed to heat, using its own internal exothermic reactions, for about 2 to 3 days. The fibre is then pressed in spindle presses to recover second grade (technical) oil that is used normally in soap making. The nuts are usually dried and sold to other operators who process them into palm kernel oil.

Properties of palm biodiesel oil

Generally the higher oil's Iodine value, the lower the temperature at which it solidifies. Different terms are used for this – Melting point (MP), Cloud point (CP), cold filter plugging point (CFPP) and pour point (PP). In practice they all mean about the same. It matters with both SVO systems using straight vegetable oils as fuel and with biodiesel but moiré so with the SVO systems.

Property	Unit	EN 14214 [a]	Palm Biodiesel	
			Normal Grade	Winter Grade
Ester content	% weight	Min. 96.5	> 98.0	> 98.0
Density @ 15°C	kg/m³	860 - 900	878	870 - 890
Viscosity @ 40°C	mm²/s	3.5 – 5.0	4.4	4.0 - 5.0
Flash point	°C	Min. 120	182	150 - 200
Pour point	°C	--	15	-21 to 0
CFPP	°C	See note (b)	15	-18 to - 3
Iodine value	--	Max. 120	50 - 55	56 - 83
Cetane number	--	Min. 51	58 - 59	53 - 59
Sulphur	mg/kg	Max. 10	< 10	< 10
Phosphorous	mg/kg	Max. 10	< 10	< 10
Water	mg/kg	Max. 500	< 500	< 500
Linolenic acid methyl esters	% weight	Max. 12	< 0.5	< 0.5
Polyunsaturated methyl esters	% weight	Max. 1	< 0.1	< 0.1
Methanol	% weight	Max. 0.2	< 0.2	< 0.2
Monoglycerides	% weight	Max. 0.8	< 0.8	< 0.8
Diglycerides	% weight	Max. 0.2	< 0.2	< 0.2
Triglycerides	% weight	Max. 0.2	< 0.2	< 0.2
Free glycerol	% weight	Max. 0.02	< 0.02	< 0.02
Total glycerol	% weight	Max. 0.25	< 0.25	< 0.25

Table. Properties of Palm Biodiesel Oil.

III. PALM OIL AS AN ALTERNATIVE FUEL FOR I. C. ENGINES

There are at least three ways to run a diesel engine on vegetable oil. (a) Mix it with petroleum diesel fuel or with a solvent or with gasoline. (b) Use the oil just as it is, usually SVO – Straight Vegetable Oil or PPO – Pure Plant Oil. (c) Covert it to biodiesel. The first two methods sound easy, but so often in practice, it's not so simple. Vegetable oil is much more viscous (thicker) than either petrol-diesel or biodiesel. The purpose of mixing or blending SVO with other fuels and solvents is to lower the viscosity to make it thinner, so that it flows more freely through the fuel system into the combustion chamber. People use various mixes, ranging from 10% SVO

and 90% petro-diesel to 90% SVO and 10% petro-diesel.

- Palm Biodiesel: - We know that petroleum diesel leaves a lot of dirt in the tank and the fuel system. Biodiesel is a good solvent. It tends to free the dirt and clean it out. A common warning is that biodiesel, especially 100% biodiesel, will damage any natural or butyl rubber parts in the fuel system. The main reaction for converting oil to biodiesel is called Transesterification. These process reacts an alcohol (like methanol) with the triglyceride oils contained in vegetable oils, animal fats or recycled greases, forming fatty acid alkyl esters (biodiesel) and glycerin. The reaction requires heat and a strong base catalyst, such as sodium hydroxide or potassium hydroxide.

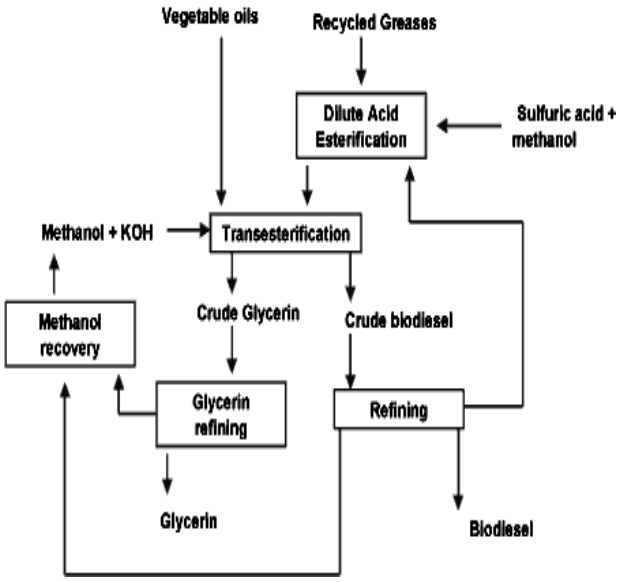


Figure. Biodiesel Production.

IV. ELSBETT CONVERSION TECHNOLOGY

To use vegetable oil as a I.C. Engine fuel, ELSBETT offers a conversion technology with precisely matched technical characteristics for almost all regular diesel engines for cars/vans, lorries and tractors. Conversion technology is needed to compensate for the special characteristics of vegetable oil such as

1. Being a product of nature, its quality varies considerable.
2. It has higher flash point than diesel.
3. It flows freely than diesel.
4. It ignites less readily.

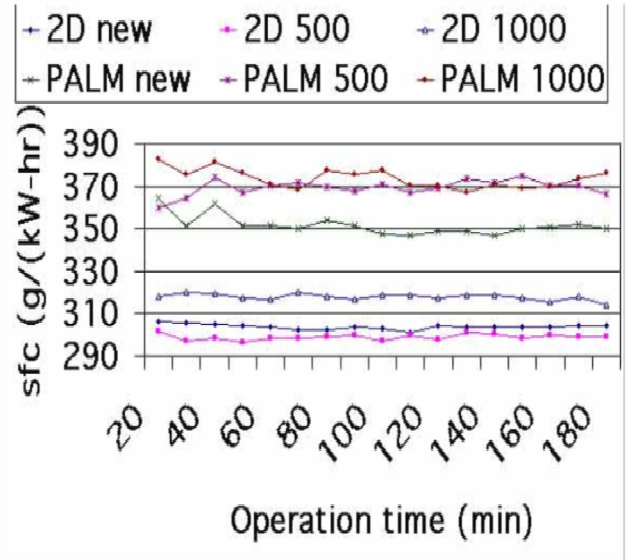
A good conversion technology will cater for the other special characteristics of vegetable oil by including the following features: - Pre-warning of the fuel, fuel lines or engine parts, modifications to the fuel system/pumps. Additional filter stages, adjustment of the electronic engine control, with the 1-tank conversion technique, possible modification to the injection system, control elements and relays.

The main issue to overcome with SVO is its viscosity. In an ELSBETT 1-tank conversion technique, this is achieved by preheating the oil and optimizing the injector spray pattern. Some of the components of this system are –

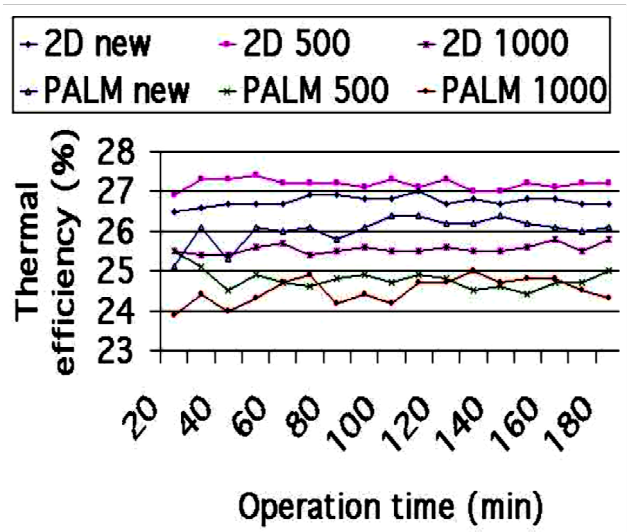
1. Injector nozzles better suited to higher viscosities.
2. Higher performance glow plugs.
3. An extra temperature switch to extend the glow plug operation.
4. An SVO fuel filter.
5. An electric heater fitted to the SVO fuel filter.
6. A coolant-to-fuel heat exchanger.

V. COMPARISON OF PALM OIL AND DIESEL

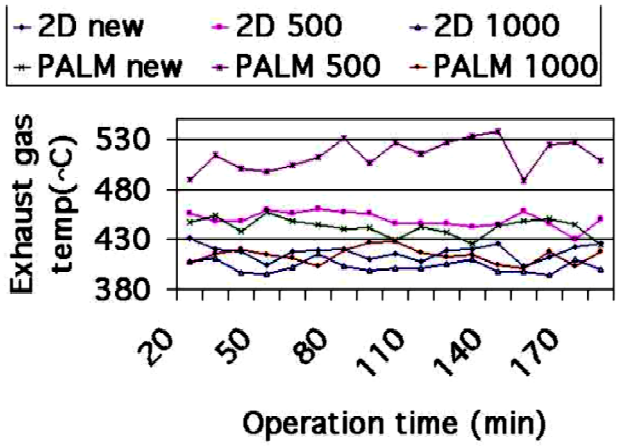
The following description deals with a comparative performance testing of diesel engine using diesel oil and refined palm oil over 2000 hours of continuous running time. Short-term performance testing was conducted for each fuel on the dynamometer engine teat bed. Specific fuel consumption, exhaust temperature and black smoke density were determined during the test. Long-term performance testing was also done by running the engines coupled with a generator in order to supply load (electricity) to a light board. For each 500 hours of engine run time, the engines were disassembled for engine wear inspection. The following graphs show the comparative performance analysis.



Graph 1. Comparison of various palm biodiesel oils with respect to SFC and Operating time.



Graph 2. Comparison of various palm biodiesel oils with respect to Thermal efficiency and Operating time.



Graph 3. Comparison of various palm biodiesel oils with respect to Exhaust Gas Temperature and Operating time.

VI. Merits and Demerits of Palm Oil

A. Merits of palm oil

- Palm oil is non toxic, biodegradable.
- It reduces the emission of harmful pollutants (mainly particulates) from diesel engines but emissions of nitrogen oxides are increased.
- Palm oil has high Cetane number. This higher value of Cetane number contributes to easy cold starting and low idle noise.
- The use of palm oil based biodiesel can extent the life of diesel engines because it is more lubricating and further more power output are relatively unaffected by biodiesel.
- Palm oil is also a safer fuel than diesel due to its high flash point.

B. Demerits of palm oil

- Palm oil production has been documented as a cause of substantial and often irreversible damage to the natural environment. Its impacts include; deforestation, habitat loss of critically endangered species and a significant increase in green house gas emissions especially NO_x.
- Environmental groups such as Greenpeace claim that the deforestation caused by making way for oil plantations is far more damaging for the climate than the benefits gained by switching to bio fuels.
- Much of the recent investment in new palm plantations fro biofuel has been part funded through the carbon credits projects through the Clean Development Mechanism.

VII. Conclusion

With the depletion of fossil fuels and the current worldwide efforts to address the issue of global warming, palm oil biodiesel has a definite advantage over the conventional diesel as palm biodiesel is made from renewable resources and is more environmental friendly. Results show that using 100% refined palm oil in a KUBOTA diesel engine and operating continuously at constant 75% maximum load and 2200 rpm for 2000 hours produce no serious problem. The black smoke density is not significantly different in the two cases. Wear in the engines fueled by refined palm oil and diesel oil is not significantly different. Hence there is not much trouble in using the palm oil instead of diesel oil. In fact the diesel engine will become somewhat green engine by using palm oil with respect to air pollution.

References

[1] Ataga, D.O. Ilechie, C. O. and Omoti, U. “Small scale palm oil processing technology in Nigeria”, Accra, Ghana, 1993.
[2] Harris, R. “A survey of Palm oil and Palm kernel extraction equipment”, technical report on FAO project, April 1995.
[3] Poku K. “Palm oil based mill project technical evaluation of machinery and equipments”, 1993.
[4] Poku K. “Innovations in village based palm oil processing”, Accra, Ghana, 1993.
[5] Poku K. “Oil palm small holder development: Processing technology”, a mission report on FAO project, 1998.
[6] Tang, S. T. and Teoh, P. K., “Palm kernel oil extraction”, world conference on Processing of palm, palm kernel; and coconut oils, Vol 62 pg 255, Feb 1985.

Modeling High Power CO_2 Laser Machining Centre with using Autodesk Inventor.

Dhaval P. Patel,
Asst. Prof., Mechanical Engineering Dept.,
Gandhinagar Institute Of Technology

***Abstract*—Autodesk Inventor is 3D mechanical design system built with adaptive technology and solid modeling capabilities. In these paper author empathies maximum capabilities of Autodesk Inventor which useful for new product design and development. Modeling of High Power CO_2 Laser Machining Centre in which designing of X, Y and Z axes of machine by parametric modeling and using all feature of Autodesk Inventor.**

***Index Terms*—Laser Cutting Machine, Autodesk Inventor**

I. INTRODUCTION

The Autodesk Inventor software include features for 3D modeling, information management, collaboration, and technical support with Autodesk Inventor, you can:

- Create 3D models and 2D manufacturing drawings.
- Create adaptive features, parts, and subassemblies.
- Manage thousands of parts and large assemblies.
- Use third-party applications, with an Application Program Interface (API).
- Use VBA to access the Autodesk Inventor API. Create programs to automate repetitive tasks. On the Help menu, choose Programmer Help.
- Import SAT, STEP, and AutoCAD and Autodesk Mechanical Desktop (DWG) files for use in Autodesk Inventor. Export Autodesk Inventor file to AutoCAD, Autodesk Mechanical Desktop, and IGES formats.
- Collaborate with multiple designers in the modeling process.
- Link to web tools to access industry resources, share data, and communicate with colleagues.
- Use the integrated Design Support System (DSS) for help as you work.

II. CAPABILITIES OF AUTODESK INVENTOR

1. Autodesk Inventor tools

There are main four tools of Autodesk Inventor series 9

- Part modelling
- Assembly modelling
- Presentation
- Drawing

1.1 Part modelling features

1.1.1 2D to 3D features

⇒ *To create extrude feature*

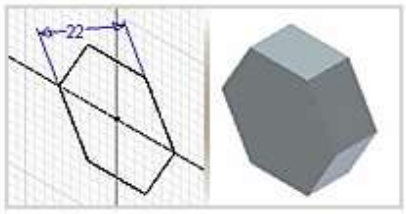


Figure 1 Extruded view of hexagonal part

⇒ *To create revolved feature*

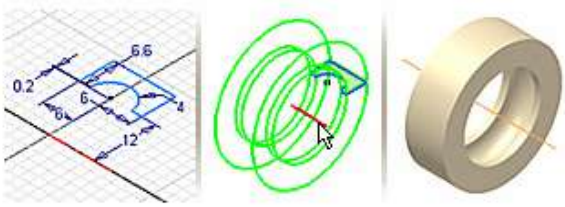


Figure 2 Revolved view of bearing

⇒ *To create sweep feature*

Figure 3 Sweep view of stepped part

1.1.2 Create features

⇒ *To fillet an edge*

Figure 4 Filleting edges of angle part

⇒ *To arrange features in a rectangular pattern*

Figure 5 Rectangle pattern of rectangle block on part

⇒ *To arrange features in a circular pattern*

Figure 6 Circular pattern of cylinder on part

⇒ *To create a mirrored feature*

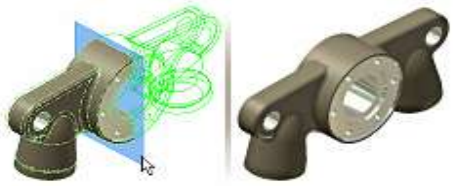


Figure 7 Mirrored of single part

⇒ *To create a hole feature*

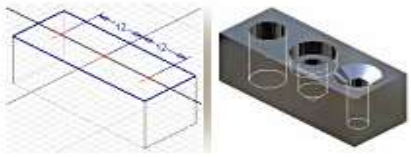


Figure 8 Different holes on rectangle block

2. Working with assemblies

- Turn off visibility of nonessential components. Access the parts we need and update graphics faster.
- Use design representations. Create design representations that highlight specific design problems or assembly subsystems, and apply them when opening the assembly model.
- Turn off part adaptivity. After we size components, turn off adaptivity to speed up solutions and prevent accidental changes.
- Assign different colors to components. Select colors from the Color list on the Standard toolbar.
- Use the browser to find components. Point to component in the browser to highlight in the graphics window.
- Use color to identify components groups. Using attributes, find components in specific subsystems or from specific vendors and color-code them in named representations.

Using part features as explained in previous section creates all components of High Power CO_2 Laser Machining Centre. All assemblies are created using various components (part) by constrained there relative motion. In High Power CO_2 Laser Machining Centre, there are main three types of sub assemblies, which are as follows:

⇒ Assembly of X – axis

⇒ Assembly of Y – axis

⇒ Assembly of Z – axis

Final model of High Power CO_2 Laser Machining Centre is created by assembling above all sub assemblies. All sub assemblies and assembly is shown in the following section.

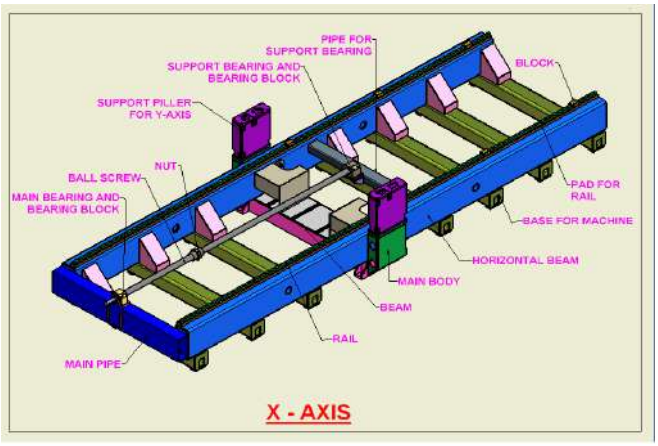


Figure 9 Assembly model of x-axis

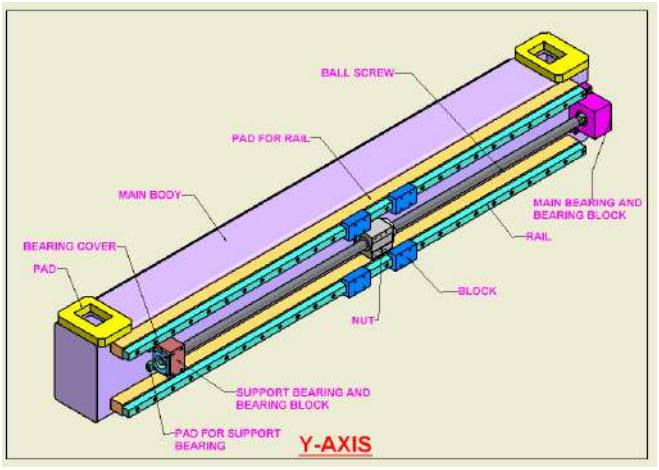


Figure 10 Assembly model of y-axis

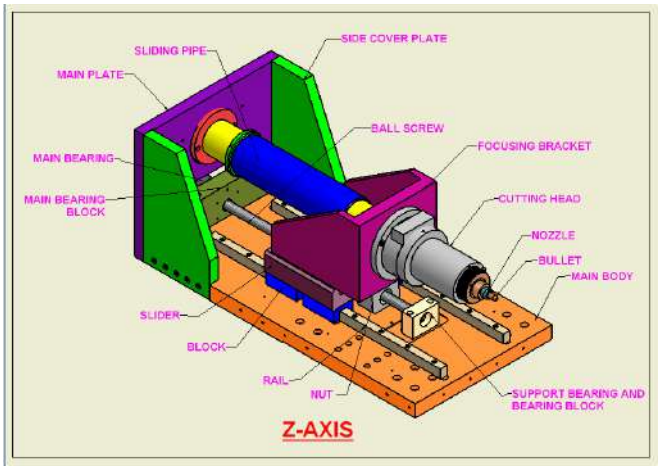


Figure11 Assembly model of z-axis

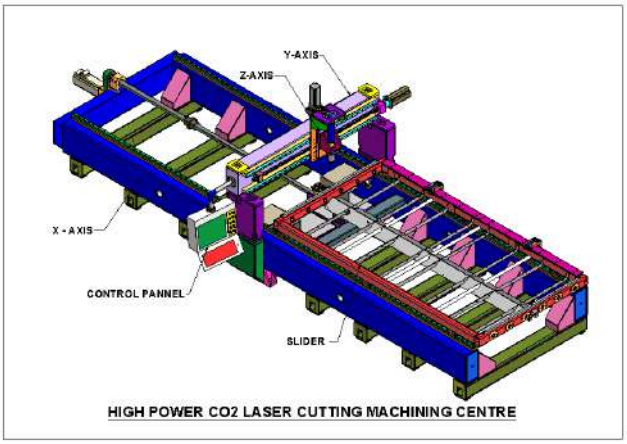


Figure 12 Assembly model of High Power Co2 Laser Machining Centre

3. Presentation (Animation)

We can develop exploded views, animations, and other stylized views of an assembly to help document your design. Any static presentation view can be used in a drawing.

Presentation views are saved in a separate file called a presentation file (.ipn). Each presentation file can contain as many presentation views as needed for a specified assembly. When changes are made to an assembly, the presentation views are updated automatically.

3.1 Develop an animation

After setting up an exploded view, you can view, edit, and fine-tune it to develop the desired animation.

- Set the browser view to Sequence.
- Select a task or sequence in the active view, right-click, and select Edit to open the Edit Task and Sequences dialog box.
- In the dialog box select a task or sequence, then:
  - ✓ Click the corresponding play button to view its animation.
  - ✓ Enter a description for it in the edit box.
  - ✓ Set the Interval to control the playback speed for the animation.
  - ✓ Use the view tools on the standard tool bar to change the orientation of the view, then click Set Camera save a camera view for the selected task.

Note: To save your changes, you must click the Apply button before you edit another task or sequence.

- To change the order of the animation, close the dialog box, and reorder the sequences and tweaks in the browser.

3.2 Record an animation

After developing an animation, you can play it when the presentation file is open or record it to an external file and replay it later. Use the Animate button on the Presentation View Management toolbar.

1. Click the Animate button.
2. Set the number of repetitions.
3. Click one of the play buttons to view the animation in the graphics window.
4. To record the animation to a file, click the record button, then click one of the play buttons start recording.

Note: To change the number of Repetitions, click the Reset button on the dialog box, then enter the new value.

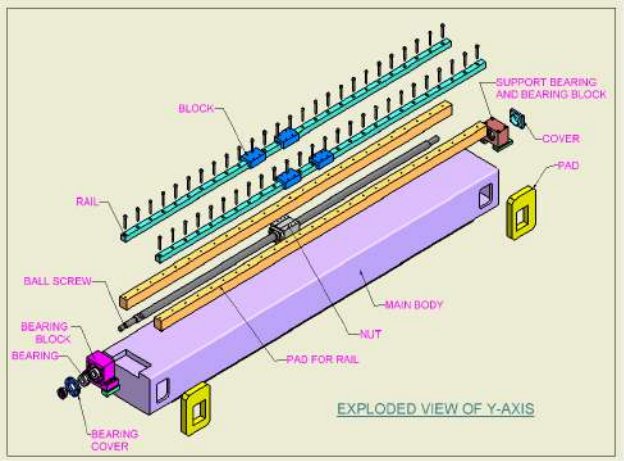


Figure 13 Exploded view of y-axis

4.Drawing (2-Dimension view)

After we create a model, we can create a drawing file (with .idw extension) to document your design. In a drawing file, we place views of a part model that can include any combination of model dimensions (parametric dimensions extracted from the part file) and drawing dimension (created in the drawing file). We can add suppress dimensions in each view as needed, and place dimensions, annotations, and symbols that adhere to styles associated with the ANSI, BSI, DIN, GB, ISO and JIS standard. We can also define own custom style.

We can change the alignment, label, scale, and displayed dimensions in any view. We can also edit our part by changing the parametric model dimensions from within the drawing file will automatically update with any change saved in part file.

4.1 Create drawing

1. Click the new button on standard toolbar, and then choose a drawing template from the Default, English or Metric tab.

   We open the template based on the template we specified when we installed Autodesk Inventor. The default drawing is a blank sheet with a border and title block. The English and Metric tabs contain the templates for those units of measure.

2. On the drawing views panel bar, click base view.

3. In the drawing view dialog box, click the explore directories button beside the file box to locate a part or assembly. If we already have a model open, it is used by default for the view.

4. Accept the default scale, label, and other settings. A preview of the view is attached to the cursor. Click on the drawing sheet to place the view and close the dialog box.

Autodesk Inventor maintains links between components and drawings, so we can create a drawing at any time during the creation of a component. By default, the drawing updates automatically when we edit the component.

We can modify the drawing border and title block to comply with our company specifications. The first folder at the top of browser is drawing resources. We can expand drawing resources to show the sheet formats, borders, title blocks, and sketched symbols that available to use in the drawing. We can customize, add to or delete items from drawing resources.

While drawing view define the shape of component, drawing annotation provides the additional information needed to complete documentation of the component. In Autodesk Inventor, style defines annotations, according to the active drawing standard. Each standard has a default set of available styles, which can be customized as needed. Custom drawing elements, such as title blocks, borders, and other standard annotations can be added to added to a drawing template so they are available in every drawing based on the template.

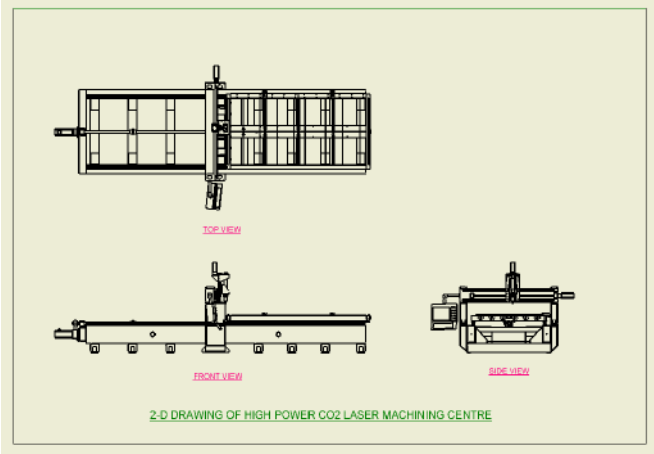


Figure14 2-D view of High Power Co2 Laser Machining Centre

4. Styles

4.1 Styles and Standards Editor creates and edits styles

Document formatting is now managed with one, the Styles and Standards editor. The Editor sets the active standard, and provides a complete list of styles associate list of styles associated with the standard. Style types are listed in a browser, which accesses all controls for the style in an editing window.

Some styles reference other styles to complete their definition. For example, dimension styles, balloon styles, feature control frames, and parts lists, use text styles to format text. Changes to a sub style update the style, and automatically update all other styles that reference it.

1) Style Libraries contain style definitions

A style library contains definitions of individual style types, so that all design projects using the library have access to the same styles. Styles assure that documents use uniform formatting and are updated easily when a style change is required.

2) Style Management wizard

A tool for administrators assists you in batch-style management for Inventor files. Using the style Management Wizard, you can:

- Harvest styles from AutoDesk Inventor files and place them in a target style library.
- Purge styles from AutoDesk Inventor files. Purging removes unused style information and can be beneficial in large assemblies by reducing required memory.

3) Purge unused styles

Individual documents may contain styles that are not used. We can purge unused styles and substyles from a document.

The Purge Styles tool automatically ignores styles that are in use.

4) Update styles with latest versions

We can easily update styles in the current document to match styles in style library.

5) Import and export styles

We can use import and export to copy styles between documents. The style is exported in a special *. styxml file, which is then imported into another AutoDesk Inventor document.

6) Layers manage display properties

AutoDesk Inventor now has layers. The Layer Style specifies a layer name and attributes for common display properties of drawing objects. Individual layers can be turned on and off and created and assigned to attributes as desired.

5. Projects

Project files are improved to make them easier to create and use. We can:

- Specify Frequently Used Subfolders so you can quickly locate files. The specified folders are added to locations when you open a file.
- Set an option called Using File Names so that you can automatically find files after they are moved to a different location.
- Specify a location for project templates.
- Specify a location for style libraries.
- Specify the location of the Content Center.

6. Data Translation

6.1 DWG compatibility

DWG export now supports:

- AutoDesk Inventor layers are exported one-to-one to AutoCAD and Mechanical Desktop layers
- Dimension and text format and size is retained.
- Title blocks and borders retain their attributes.
- OLE objects, such as logos (bitmaps) and Excel spreadsheets.

DWG import now supports:

- Layers are imported intact to AutoDesk Inventor.
- Title blocks, borders, and text are preserved.
- Performance is improved.

6.2 IGES/STEP Enhancements

We can now use imported wire frames and edges of imported solids and surface models for creating sketches. Data is imported into the construction environment, but geometry may be selected from a sketch.

III. CONCLUSION

Autodesk Inventor capabilities useful for making any mechanical system for physical appearance of object and calculate all properties such like mass, weight, area, volume and moment of inertia. Also do F.E.A. analysis on base of problem static, dynamic, thermal and CFD.In this paper thousand of component of CO_2 laser cutting machining centre are modeling by Autodesk Inventor.

REFERENCES

[1] www.sahajanandlaser.com
[2] Autodesk Inventor software‘s help documents.
[3] www.microconcepts.co.uk
[4] Dr. Sadhusingh., “Computer Aided Design and Computer Aided Manufacturing”, Third Edition: 2007.

PROJECT SCHEDULING WITH PERT/CPM TECHNIQUES

PRIYANK V PAREKH ASSISTANT PROFESSOR-MED-GIT

ABSTRACT-Program evaluation and review technique (PERT) and the critical path method (CPM) are used to plan, coordinate, and control the performance of activities that comprise complex projects. Designed as a tool for executing projects that are unique and non-repetitive, PERT methods typically include the preparation of a network diagram. Constructing a PERT diagram requires a manager to identify the interdependent phases of a project and the correct sequences of activities that lead to completion. Furthermore, when combined with applications of probability theory, PERT may be used to predict the duration of activities and forecast the complete date of a project Closely related to PERT, CPM techniques are used to guide project managers in making resource allocation decisions during the life of a project. Critical path method guides managers in identifying the sequence of activities that most directly influence the length of a project. Moreover, CPM enhances a manager's ability to use resources in an efficient way by providing a strategy for determining which activities can be expedited to minimize total project costs. Together, PERT and CPM rationalize the central tasks of project management projects, while coordinating resource allocations to maximize efficiency and effectiveness

Keywords: Program evaluation and review technique (PERT); Critical path method (CPM); Event; Activity; Probability theory; Critical path; Project management

I INTRODUCTION:

Program Evaluation and Review Technique (PERT) and Critical Path Method (CPM) are diagrams applicable for both the planning and control aspects of production. Visual display of the network enhances the communication and highlights the interdependency of the various activities required for project completion. Perhaps the greatest contribution of these tools is the identification of sequentially time-critical activities that require the closest monitoring.

PERT is a method to analyze the tasks involved in completing a given project, especially the time needed to complete each task, and identifying the minimum time needed to complete the total project.

PERT was developed primarily to simplify the planning and scheduling of large and complex projects. It was able to incorporate uncertainty by making it possible to schedule a project while not knowing precisely the details and durations of all the activities. It is more of an event-oriented technique rather than start- and completion-oriented, and is used more in R&D-type projects where time, rather than cost, is the major factor.tools widely used in project scheduling. Both are based on network analysis.

II BACKGROUND

In the early 1900s the Gantt chart was widely hailed as the reason that ships were built in record time. Developed by an engineer named Henry Gantt, this horizontal bar chart shows the scheduled times for individual jobs to be accomplished by specific resources. However, this tool is static in nature, and requires frequent manual updating, especially when activities are sequentially dependent.

In Figure 1, the Gantt chart shows the prospective times for five activities in a project, but does not show an underlying dependency of Activity D on the completion of Activity B.

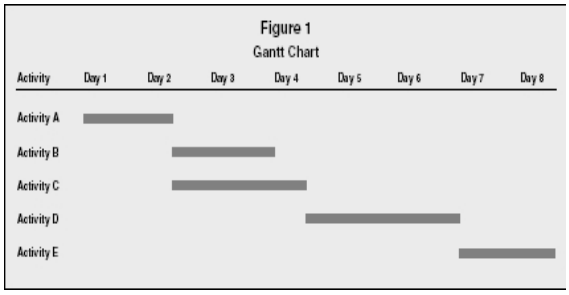


Figure 1 Gantt chart

In the 1950s, two groups independently developed what has become known as the PERT/CPM method of project scheduling. Each of these techniques improved on the Gantt chart by building into the tool the explicit sequencing of activities.

PERT was developed by the U.S. Navy, the Lockheed Corporation, and the consulting firm of Booz, Allen and Hamilton to facilitate the Polaris missile project. As time was a primary issue, this technique used statistical techniques to assess the probability of finishing the project within a given period of time.

By contrast, CPM was created in the environment of industrial projects, where costs were a major factor. In addition to the identification of the time-critical path of activities, representatives from the Du Pont Company and Sperry-Rand Corporation also developed a time-cost tradeoff analysis mechanism called crashing. These two tools differ in the network diagram display. PERT historically uses the activity-on-arrow (AOA) convention, while CPM uses activity-on-node (AON). For most purposes, these two conventions are interchangeable; however some propriety software requires the logic of a specific convention. Both forms of network diagrams use arrows (lines implying direction) and nodes (circles or rectangles) to define the set of project activities or tasks. The flow of logic is from left to right.

To simplify the diagram, letters are frequently used to represent individual activities. Figures 2 and 3 illustrate the differences for the same simple project.

Figure 2 illustrates the AOA convention, in which arrows depict activity requiring time and resources. The node represents an event, which requires neither time nor resources; this event is actually recognition that prior tasks are completed and the following tasks can begin. While the length of the arrow is not necessarily related to the duration of the task, there may be a tendency on the part of the analyst to sketch longer arrows for longer activities. To maintain the integrity of the network, there may be need for a dummy activity, as it is not acceptable to have two tasks that share the same beginning and ending nodes.

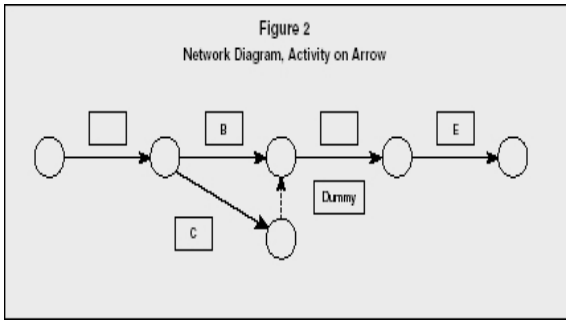


Figure 2 Activity on Arrow Diagram

III Scheduling of a Project using CPM(CRITICAL PATH METHOD):

Scheduling is an important part of the planning of any project. However, it is first necessary to develop a list of all the activities required, as listed in the work breakdown structure.

Activities require both time and the use of resources. Typically, the list of activities is compiled with duration estimates and immediate predecessors.

To illustrate the use of CPM, we can imagine a simple cookie-baking project: the recipe provides the complete statement of work, from which the work breakdown structure can be developed. The resources available for this project are two cooks and one oven with limited capacity; the raw materials are the ingredients to be used in preparing the cookie dough. As listed in Table 1, the activities take a total of 80 minutes of resource time. Because some activities can run parallel, the cooks should complete the project in less than 80 minutes.

Table 1

List of Project Activities (CPM)

Description of Activity	Duration (minutes)	Immediate Predecessor(s)
A. Preheat oven	15 minutes	—
B. Assemble, measure ingredients	8 minutes	—
C. Mix dough	2 minutes	B
D. Shape first batch	4 minutes	C
E. Bake first batch	12 minutes	A, D
F. Cool first batch	10 minutes	E
G. Shape second batch	4 minutes	C
H. Bake second batch	12 minutes	E, G
I. Cool second batch	10 minutes	H
J. Store cookies	3 minutes	F, I
Total time	80 minutes	

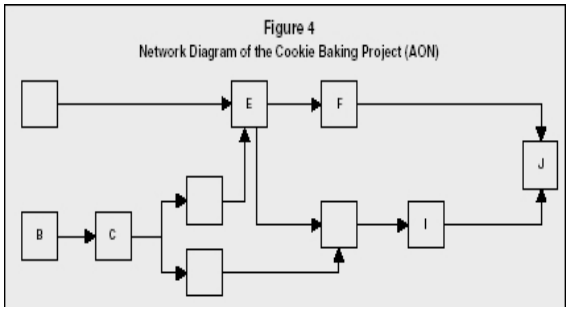


Figure 3 Network Diagram

After beginning the project at 8:00 A.M., the first batch of dough is ready to go into the oven at 8:14, but the project cannot proceed until the oven is fully heated—at 8:15. The cooks actually have a one-minute cushion, called slack time. If measuring, mixing, or shaping actually take one additional minute, this will not delay the completion time of the overall project.

Figure 3 illustrates the network diagram associated with the cookie-baking project. The set of paths through the system traces every possible route from each beginning activity to each ending activity. In this simple project, one can explicitly define all the paths through the system in minutes as follows:
A-E-F-J = 15 + 12 + 10 + 3 = 40
A-E-H-I-J = 15 + 12 + 12 + 10 + 3 = 52
B-C-D-E-F-J = 8 + 2 + 4 + 12 + 10 + 3 = 39
B-C-D-E-H-I-J = 8 + 2 + 4 + 12 + 12 + 10 + 3 = 51
B-C-G-H-I-J = 8 + 2 + 4 + 12 + 10 + 3 = 39

The managerial importance of this critical path is that any delay to the activities on this path will delay the project completion time, currently anticipated as 8:52 A.M. It is important to monitor this critical set of activities to prevent the missed due-date of the project. If the oven takes 16 minutes to heat (instead of the predicted 15 minutes), the project manager needs to anticipate how to get the project back on schedule. One suggestion is to bring in a fan (another resource) to speed the cooling process of the second batch of cookies; another is to split the storage process into first- and second-batch components.

Other paths tend to require less monitoring, as these sets of activities have slack, or a cushion, in which activities may be accelerated or delayed without penalty. Total slack for a given path is defined as the difference in the critical path time and the time for the given path. For example, the total slack for B-C-G-H-I-J is 13 minutes (52–39 minutes). And the slack for B-C-D-E-H-I-J is only one minute (52–51), making this path near critical. Since these paths share some of the critical path activities, it is obvious that the manager should look at the slack available to individual activities.

Table 2 illustrates the calculation of slack for individual activities. For projects more complex than the simplistic cookie project, this is the method used to identify the critical path, as those activities with zero slack time are critical path activities. The determination of early-start and early-finish times use a forward pass through the system to investigate how early in the project each activity could start and end, given the dependency on other activities.

Table 2

Calculation of Slack Time

Activity	Early Start	Early Finish	Late Start	Late Finish	Slack
A	8:00	8:15	8:00	8:15	0
B	8:00	8:08	8:01	8:09	1
C	8:08	8:10	8:09	8:11	1
D	8:10	8:14	8:11	8:15	1
E	8:15	8:27	8:15	8:27	0
F	8:27	8:37	8:39	8:49	12
G	8:10	8:14	8:23	8:27	13
H	8:27	8:39	8:27	8:39	0
I	8:39	8:49	8:39	8:49	0
J	8:49	8:52	8:49	8:52	0

The late-time calculations use the finish time from the forward pass (8:52 A.M.) and employ a backward pass to determine at what time each activity must start to provide each subsequent activity with sufficient time to stay on track.

Slack for the individual activities is calculated by taking the difference between the late-start and early-start times (or, alternatively, between the late-finish and early-finish times) for each activity. If the difference is zero, then there is no slack; the activity is totally defined as to its time-position in the project and must therefore be a critical path activity. For other activities, the slack defines the flexibility in start times, but only assuming that no other activity on the path is delayed

PERT(PROGRAMME EVALUATION AND REVIEW TECHNIQUE):

The main objective of PERT is to facilitate decision making and to reduce both the time and cost required to complete a project.

PERT is intended for very large-scale, one-time, non-routine, complex projects with a high degree of intertask dependency, projects which require a series of activities, some of which must be performed sequentially and others that can be performed in parallel with other activities.

PERT planning involves the following steps that are described below.

1. Identify the specific activities and milestones. The activities are the tasks required to complete a project. The milestones are the events marking the beginning and the end of one or more activities. It is helpful to list the tasks in a table that in later steps can be expanded to include information on sequence and duration.

2. Determine the proper sequence of the activities. This step may be combined with the activity identification step since the activity sequence is evident for some tasks. Other tasks may require more analysis to determine the exact order in which they must be performed.

3. Construct a network diagram. Using the activity sequence information, a network diagram can be drawn showing the sequence of the serial and parallel activities. Each activity represents a node in the network, and the arrows represent the relation between activities. Software packages simplify this step by automatically converting tabular activity information into a network diagram.

4. Estimate the time required for each activity. Weeks are a commonly used unit of time for activity completion, but any consistent unit of time can be used. A distinguishing feature of PERT is its ability to deal with uncertainty in activity completion time.

For each activity, the model usually includes three time estimates:

- **Optimistic time** – Generally the shortest time in which the activity can be completed. It is common practice to specify optimistic time to be three standards deviations from the mean so that there is a approximately a 1% chance that the activity will be completed within the optimistic time.

- **Most likely time** – The completion time having the highest probability. Note that this time is different from the *expected time.*

- **Pessimistic time** – the longest time that an activity might require. Three standard deviations from the mean is commonly used for the pessimistic time.

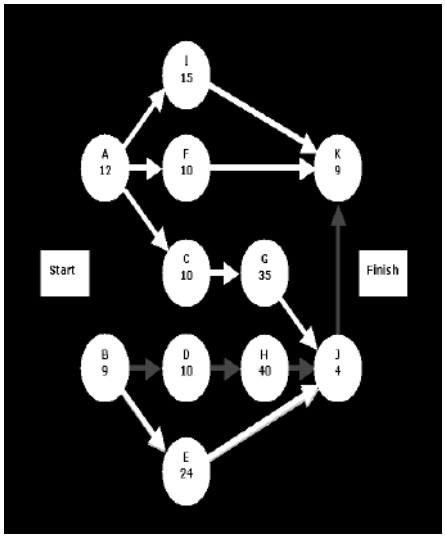


Figure 4 PERT Network

PERT assumes a beta probability distribution for the time estimates. For a beta distribution, the expected time for each activity can be approximated using the following weighted average:

Expected time = (Optimistic + 4 x Most likely + Pessimistic) / 6

5. Determine the critical path. The critical path is determined by adding the times for the activities in each sequence and determining the longest path in the project. The critical path determines the total calendar time required for the project. If activities outside the critical path speed up to slow down (within limits), the total project time does not change. The amount of time that a non – critical path activity can be delayed without the project is referred to as a *slack time.* If the critical path is not immediately obvious, it may be helpful to determine the following four quantities foe each activity: ES – Earliest Start time EF - Earliest Finish time LS – Latest Start time LF - Latest Finish tim

Table 3 PERT data

ACTIVITY	DESCRIPTION	PRECEDANCE	OPTIMISTIC TIME	MOST LIKELY TIME	PESSIMISTIC TIME	EXPECTED TIME
A	Select Administrative and Medical staff		9	12	15	12
B	Select site And do site survey		5	9	13	9
C	Select Equipment	A	8	10	12	10
D	Prepare plans And layouts	B	7	9	17	10
E	Bring utilities To sites	B	18	23	34	24
F	Fill the staff position	A	9	9	15	10
G	Purchase and Take delivery of Equipment	C	30	35	40	35
H	Construct the hospital	D	35	39	49	40
I	Develop an information system	A	12	15	18	15
J	Install the equipment	E,G, H	3	3	9	4
K	Train nurses and support staff	F,I,J	7	9	11	9

The earliest start time, earliest finish time, latest start time and latest finish time for each activity are calculated in the following table:

Table 4 PERT Calculations

NODE	DURATION	ES	EF	LS	LF	SLACK
A	12	0	12	2	14	2
B	9	0	9	0	9	0
C	10	12	22	14	24	2
D	10	9	19	9	19	0
E	24	9	33	35	59	26
F	10	12	22	53	63	41
G	35	22	57	24	59	2
H	40	19	59	19	59	0
I	15	12	27	48	63	36
J	4	59	63	59	63	0
K	6	63	72	63	72	0

REFERANCES

1. Archibald & Villoria, "Network-Based Management Systems (PERT/CPM)," John Wiley & Sons, 1967
2. Cagle, Ronald B. Blueprint for project recovery: a project management guide : the complete process for getting derailed projects back on the track. New York : AMACOM, 2002
3. Clet, Étienne and Henri-Pierre Maders. Comment manager un projet. 2nd ed. Paris: Éditions d'Organisation, 2005.
4. Denis Lock, *Managementul de proiect*. Ed. CODECS, Bucureşti, 2000.
5. DeWitte, "Manpower Leveling of PERT Networks," Data Processing for Science/Engineering, March-April 1964 .
6. Frame, J. Davidson. Managing projects in organizations : how to make the best use of time, techniques, and people. San Francisco : Jossey Bass, 1995.
7. Frame, J. Davidson. The new project management : tools for an age of rapid change, complexity, and other business realities. San Francisco : John Wiley & Sons, Inc., 2002.
8. Heerkens, Gary R.Project management. New York : McGraw-Hill, 2002
9. Heldman, Kim. Project management jumpstart. San Francisco : Sybex Books, 2003.
10. Kendrick, Tom. Identifying and managing project risk : essential tools for failure-proofing your project. New York : AMACOM Books, 2003.
11. Levine, Harvey A. Practical project management : tips, tactics, and tools. New York : John Wiley & Sons Inc., 2002.
12. Lewis, James P. Project planning, scheduling and control : a hands-on guide to bringing projects in on time on budget. 3rd ed. Toronto : McGraw-Hill, 2001.
13. Phillips, "Fifteen Key Features of Computer Programs for CPM and PERT," Journal Industrial Engineering, January-February 1964

Filled PTFEs

Success Glory for Wear Parts of Reciprocating Gas Compressors over Metals & Alloys

Prof Kedar Bhojak[a], Prof Durgesh Jani[a], Prof Bhoomi Trivedi[a]
[a]Indus Institute of Technology & Engineering, Ahmedabad-382 115.

Abstract— The use of filled PTFE (Polytetraflouroethylene, Teflon) in making of the wear parts like rider rings, piston rings and packing rings for the reciprocating gas compression system became common practice since material "r-evolution" age, i.e. 1970s. Some of the industries faced the problem of high wear and tear of the piston rings, rider rings, packing rings, cooling rings etc. made up of metallic materials like bronze, cast iron, steel and alloys, k-monel, inconel etc. in conjunction of the use of different gases as the working substances. Also, as per the requirements of the end user, industry needs to manufacture the compressor parts for non lubricated, high speed, with the application environment of the high temperature and the high pressure, and hazardous gas as the working compressible fluid. Also, while finding the applications in the reciprocating gas compressors, one should consider the load on the piston rod due to the reciprocating mass and unbalanced forces and their moments at the small end of the connecting rod. So, to reduce these effects one should go for the material which possesses high wear and tear resistance along with retaining the properties and characteristics at high temperatures and pressures keeping light weight requirement in mind. The loss of life of the designed products encouraged the industry to produce the composites that can be used very efficiently in these applications. The products they incorporated in to the market are Proprietary fillers in PTFE; Carbon fiber/PTFE PEEK based, Graphite/PTFE-filled polyimide, Premium PPS/PTFE alloy, Proprietary PEEK based, Carbon graphite-filled PTFE, Glass-filled PTFE, Phenolic/cotton laminate "Micarta", Bronze-filled PTFE. The use of the above mentioned materials in making of the products such as piston rings, rider rings, cooling packing case rings etc. for reciprocating gas compressors has become the common practice nowadays. At the end of the paper the case study has been done to prove filled PTFEs, the best replacement substitutes for metallic wear parts like piston rings, rider rings and packing rings in the reciprocating gas compression systems in all aspects.

Index Terms: Composite material of wear parts, Filled PTFEs, Reciprocating Compressors.

I. INTRODUCTION

As shown in to the figure[1] below there are two major parts of the reciprocating gas compressors:

1) Frame and Running Gear Assembly
2) Gas End Assembly

Frame and Running Gear Assembly comprises of the main parts like crank shaft, connecting rod, cross head, piston rod, distance pieces, packing rings etc., while Gas End Assembly comprises of the main parts like cylinder, valves, piston, piston and rider rings,

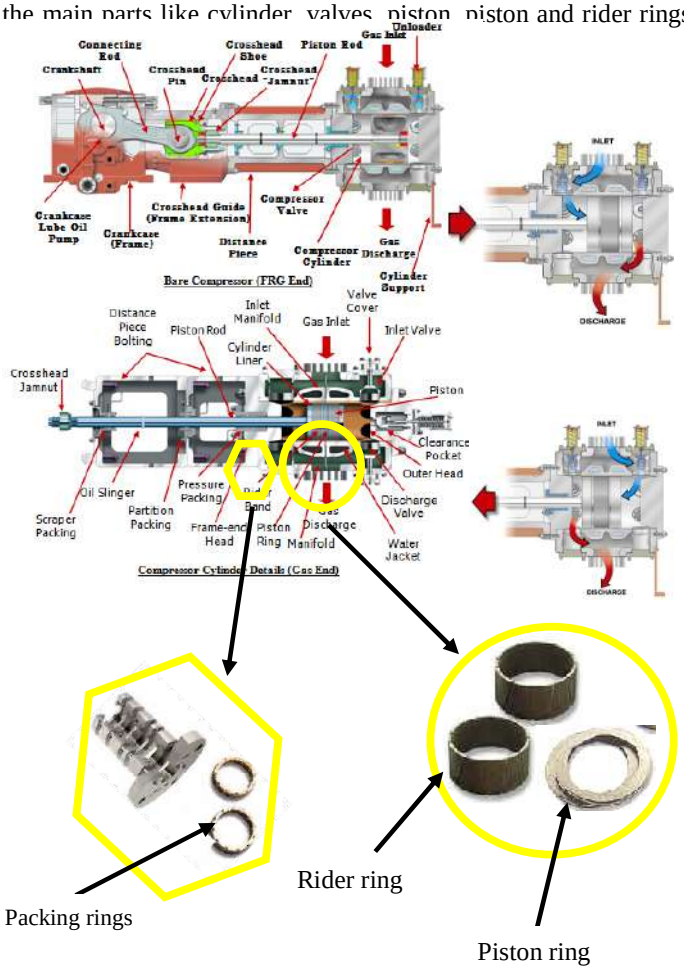
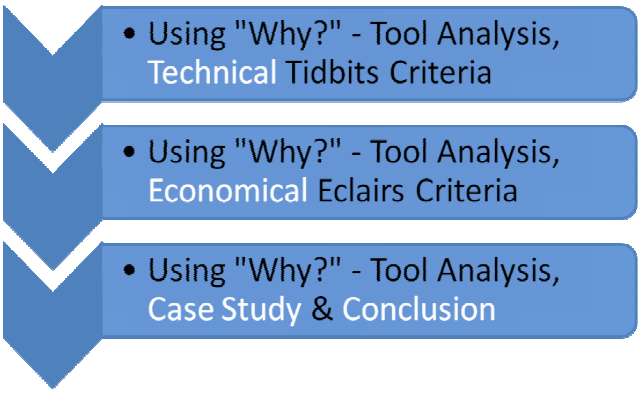


Fig 1: Cross section of Gas Compressors with Wear Parts
During this paper we will see how the selection of the different material for different applications can be done, what are the companies developed these products especially for their use, who are the manufacturers of these material parts, properties of the different materials and advantages achieved. Also, to prove the betterment of the filled PTFEs over the conventional metallic materials this in this paper I specially used the "Why? Why? Why?" – Tool of analysis. The paper flows as per the flow chart given below:



II-TECHNICAL TIDBITS CRITERIA

We are much concerned regarding to the use of appropriate materials for Packing Rings, Piston Rings and Rider Rings because, they are the most critical components in the system since they are most wearing parts, designed for small life as compared to other parts along with limited shape selection and strict tolerance in quantity of the usage. Also, their functions are very critical and important in the compression systems. So, for designing these parts along with reduction in weight and higher strength the only way left is selection of the appropriate materials.

Before the 1970s, there were very less materials available to select from, for these wear parts because of the only availability of alloys, pure metals, polymers and ceramics. After 1970s the composites came into the market and their increased usage inspired industry to use them for these wear parts. Before we discuss the current materials, I would like to discuss the required properties for materials to be used for these parts and the suitability as per the Lube and Non Lube applications and properties of conventional materials. Also, I would like to separate these different selection criteria of the materials for these wear parts.

Criteria # 1
Balancing of Reciprocating Gas Compressor and Selection of Materials for Wear Parts:

For a simple piston crank mechanism shown in figure [2], there are two types of inertial forces: Centrifugal force and reciprocating force.

The centrifugal force is generated by the reciprocating mass whose center of gravity is eccentric to center of rotation. The expression is: F= mrw2, the direction of the force is always radially outward from center of gravity of the rotating mass. This force has only one frequency i.e. circular frequency or rotation speed.

The reciprocating force is generated by the back and forth motion of the reciprocating mass. It acts in a direction in the line with piston travel. Due to non-uniform rate of acceleration and deceleration, a Fourier equation is required to determine the total force; i.e.

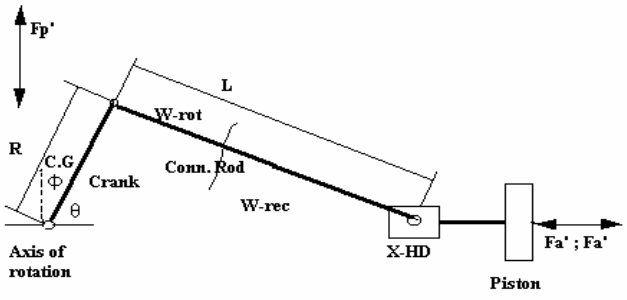
$$F = mrw2 (\cos\theta + A \cos2\theta + B \cos4\theta + C \cos6\theta + \ldots\ldots)$$

The frequency of the primary term is w or machine speed, frequency of the 2nd term is 2w or twice the machine speed and so on.
Accepted practice is to use the ratio of crank radius to connecting rod length (R/L) for the "A" constant. The 4th or higher harmonics for practical purposes can be ignored as value of "B", "C", etc diminish very rapidly. So the expression for reciprocating forces reduces to:

$$F = mrw2 (\cos\theta + R/L \cos2\theta)$$

As only magnitude of the forces is required, the periodic part can be taken as $\cos\theta=1$ i.e. $\theta=0$;However, when the angle to the center of gravity of the rotating mass does not coincide with the crank angle θ , a correction must be made for the angle ϕ by which the center gravity of the rotating mass lags or leads θ.

Fig 2: Simple piston crank system



Simple Piston-Crank System

In the figure Fp' is the primary force normal to the piston travel. There is no Fp", as the rotating force has no secondary component. Also Fa' is the primary force and Fa" is the secondary force in line with, piston travel.

Primary horizontal force, Fa' = KRN2 (W-rec + W-rot) [ϕ=0 or 180]
Secondary horizontal force, Fa" = KR(R/L) N2 W-rec
Primary vertical force, Fp' = KRN2 W-rot

For Multiple cylinder machines all the forces from individual cylinders are vectorially combined to determine the total unbalance.

Standard methods for reducing unbalance:

There are six methods reduce unbalanced forces or moments. Judgment is required when using these methods, as some times a particular method will reduce one type of unbalance at the expense of increasing another type.

Equalizing weights:
Counter weights:
Counter-weights elsewhere on the Shafting:
Removing excess weight:
Dummy Crossheads or Bob-weights:

Special balancing devices:

Efforts should be made in balancing horizontal unbalance forces. To reduce the cost and efforts made to balance the reciprocating gas compressors one should design the system such that, they have the minimum possible reciprocating mass. For this main reason the industry had developed the specialized materials like Filled PTFEs, for the reciprocating parts and the specially wear parts. They have less density for the same volume of the parts as compared to the parts made up of metallic materials.

According to the current technology the main reciprocating parts are Cross Head, Piston Rod, Piston, Packing Rings, Rider Rings, Piston Rings, and Jam Nut etc. Industry had changed the materials and shapes of the parts over existing materials to achieve this criterion.

Criteria # 2

As Per Required Properties of the Materials for Piston Rings, Rider Rings and Packing Rings:

High Wear Resistance
Low Weight
Low Friction
Excellent Chemical Resistance
High Creep Resistance
High Load and Bend Resistance
High Heat Dissipation
Electrically Inactive
No Dielectric Properties

Based on the above properties and the availability constrains the industry was using the metals and alloys for making the wear parts. These materials were the Bronze, Cast Iron, K-Monel and Inconel – White Metal Coated. But after the composites came into the market, they replaced these conventional metals. These materials are Filled PTFEs, like Carbon Filled PTFE, Glass Filled PTFE, and Molybdenum Filled PTFE etc. So, as per this criterion also, the newer composites are more appropriate than previously used materials.

Criteria # 3

As Per Required Application of Compression System like Lube and Non-Lube:

As per the requirement of the customers, companies have to provide the compression systems for bone-dry applications. Major part of the clients comprises the Petro-Chemical Industry. This particular industry has dry-atmosphere requirements due to the fire hazards studies and API-618 design criteria. They have bone-dry hydrogen substance as the working fluid most of the times. For these kinds of working substances the non-lube system is most suitable. For non-lube systems the metallic wear parts show much lesser life than the composites of PTFE do, due to their self-lubricating properties. They have shown 3 to 40 times better life than previously used metallic materials. The various advantages and better properties of the filled PTFEs over Metallic materials are as per given below:

More forgiving on mating components
Ability to run non-lube
Insurance on lubed compressors if lubrication malfunctions
Conforms better in cylinders that are not round

Tables Presentation of Filled PTFEs versus Metals & Alloys for Wear Parts as an Essence of all three Technical Tidbits Criteria

Table 1: Manufacturing Properties Comparison

PTFE+ Fillers	COTE - Radial	COTE - Axial	Max Temp. in 'C	Density g/cc	Shore D Hardness	Best Versus
C22% Graphite	55	81	250	2.02	67	K-Monel
Glass29%	80	118	200	2.21	60	Al, Bronze
Glass-Moly	104	153	200	2.26	60	-
C35%	61	90	250	2.07	62	K-Monel
C25%	68	103	200	2.03	62	-
C48%	48	71	250	2.09	65	K-Monel
C49%	23	33	250	1.95	65	Al, Bronze
C52%	97	135	200	3.8	70	Al, Bronze
C54%	73	107	200	2.07	64	Al, Bronze, K-Monel
C57%	78	115	200	2.2	63	-
C79%	62	132	200	3.94	66	-
C50%-Moly	54	108	250	1.86	69	Al, Bronze, K-Monel
C-Moly-PEEK	27	65	250	1.44	83	Al, Bronze, K-Monel
C-MOLY-PTFE	88	113	200	2	64	Al, Bronze, K-Monel

Table 2: Application & Service Comparison

Material	Application	Service	Best Versus
C22%-Graphite-PTFE	Air, NH3, CH4	NL, L	K-Monel
Glass29% -PTFE	CO, O2	NL	Al, Bronze
Glass-Moly -PTFE	-	-	-
C35%-PTFE	Air	NL	K-Monel
C25%-PTFE	-	-	-
C48%-PTFE	Air	NL	K-Monel
C49%-PTFE	O2, Cryogenic	NL	Al, Bronze
C52%-PTFE	O2, Air	NL	Al, Bronze
C54%-PTFE	H2, N2, Air, CH4	NL	Al, Bronze, K-Monel
C57%-PTFE	-	-	-
C79%-PTFE	-	-	-
C50%-Moly-PTFE	CO2, H2, C2H2	NL	Al, Bronze, K-Monel
C-Moly-PEEK	H2, NH3, CO2	NL	Al, Bronze, K-Monel
C-MOLY-PTFE	Natural Gas	L	Al, Bronze, K-Monel

Table 3: Intrinsic Material properties comparison

PTFE+Fillers	T.S. MPa	Elongation %	Disch. Pressure in psi	Prev. Life in Hrs.	Life Achieved in Hrs.	Best Versus
C22%-Graphite	13.9	12.1	1500	4500	10000	K-Monel
Glass29 %	18.7	238	-	-	-	Al, Bronze
Glass-Moly	23.3	195	-	-	-	-
C35%	12.5	19.4	-	-	-	K-Monel
C25%	12.3	75.2	-	-	-	-
C48%	12	33.4	-	-	-	K-Monel
C49%	7.3	4	-	-	-	Al, Bronze
C52%	16	101	-	-	-	Al, Bronze
C54%	15.2	86.4	1030	10000	24000	Al, Bronze, K-Monel
C57%	15.8	208	-	-	-	-
C79%	16	28	-	-	-	-
C50%-Moly	12	4	580	12000	24000	Al, Bronze, K-Monel
C-Moly-PEEK	103	3.7	510	8000	16000	Al, Bronze, K-Monel
C-MOLY	12.1	48.1	-	-	-	Al, Bronze, K-Monel

PTFE as a Matrix Material:

Introduction:

Polytetraflouroethylene or polytetrafluoroethene (PTFE) is a synthetic fluoropolymer. PTFE is most well known by the DuPont for the given brand name as “Teflon”.

Properties:

PTFE is a fluorocarbon solid, as it has a high molecular weight and it is a compound consisting wholly of carbon and fluorine.
Fluorocarbons are not as susceptible to the Van-Der-Waals force due to the high electro negativity of fluorine. Therefore, water and water containing substances, oil and oil containing substances, do not wet PTFE, as adhesion to PTFE surfaces is inhibited. Due to this property PTFE is used in the reciprocating gas compressors where working substances have the vapor portion in the composition.
It is very non-reactive, because of the strength of carbon–fluorine bonds and that’s why it is often used in these compression systems for reactive and corrosive chemicals atmosphere.
It has self lubrication properties. Where used as a lubricant, PTFE reduces friction, wear and energy consumption of machinery.
It has low COTE and highly incorporated shape retaining properties.
It has good manufacturing properties, easy availability in verities since it is the polymer and also, easily can be produced with filler materials like Carbon, Molybdenum, Glass, PEEK etc.
It and its composites and alloys are easily machined to produce required shape with low cost conventional machining processes.
It has a very much good capacity to retain all properties at higher temperatures and has the capacity to withstand thermal shocks and dielectric shocks successfully for longer time.

Applications of PTFE & Teflon®:

PTFE is used extensively in many different industries:

+Uses in petrochemical and chemical processing:
PTFE is widely used for gaskets, vessel linings, pump interiors, washers, rings, seals, spacers, dip tubes and well-drilling components. Because, it is corrosion resistant and chemically inert; it is unaffected by virtually all acids and caustics and does function in environments up to 500 degrees of Fahrenheit.

+Uses in electrical fields applications:
PTFE is one of the best insulators very well known. In thin sections, it will insulate up to 495 approximate volts per mm. There are grades of PTFE, which have a greater dielectric strength. It is used in making of wire and cable wrap, and to separate conductive surfaces in capacitors. Thick walled close-tolerance extruded tubing is the PTFE’s shape of choice, where machining or drilling long lengths to close tolerances is impossible. Multi-hole tubing can be extruded. PTFE can be machined into standoff insulators, and many different types of high voltage encapsulation devices for electrical components.

+Uses in the semi-conductor industry fields:
PTFE is inert, and its operating temperature range is from -ve350 degrees of Fahrenheit to +ve550 degrees of Fahrenheit. When made to ultra pure standards, it is the material used for various items used in chip manufacturing, including encapsulation devices for quartz heaters etc.

+Uses in the fields of food, beverage and pharmaceutical industries:
PTFE is approved by the Food and Drug Administration for use in the food, beverage, cosmetics and pharmaceutical industries. Thin Film and sheets make an inert, no-toxic slide surface without microscopic depressions where microbes can grow. Conveyance components - profiles, guide rails and slides - can withstand high temperatures inside baking and drying ovens and other heated segments of the food, cosmetics or pharmaceuticals manufacturing processes.

+Uses in the laboratory applications:
PTFE piping, tubing and vessels are used in a laboratory wares because of their chemical resistance, and inertness. No contaminants adhere to the surface of the objects.

+Uses in the bearing pads:
PTFE filled with glass, can be cut or stamped into the bearing pads, and will resist all weather-related degradation, while remaining an inert interface between disparate construction materials, like steel and concrete.

Standard shapes, structures and sizes available in the market as per ASTM1 Specifications:

PTFE Sheet and Film - ASTM D 3293 (and MIL-P22241)

Standard of the properties and characteristics parameters for PTFE molded sheets (thicker than .250 inches).

PTFE Rod - ASTD D 1710 (and MIL-P 19468)

Standard of the properties and characteristics parameters for different grades of PTFE extruded and molded rod and for heavy-walled tubing.

PTFE Molded Sheet or Shapes - ASTM D 3294

Standard of the properties and characteristics parameters for molded PTFE shapes and molded PTFE sheets.

PTFE Skived Tape - ASTM D 3308

Standard of the properties and characteristics parameters for skived tape (.250 inches and thinner)

Filled Compounds made with PTFE - ASTM D 4745

Standard of the properties and characteristics parameters for filled molding compounds made with PTFE. The specification provides standards for bulk density, tensile strength and elongation of PTFE filled with different percentages of glass fiber, glass fiber and Molybdenum Disulfide, graphite, carbon and graphite, bronze, bronze and Molybdenum Disulfide and stainless steel.

PTFE Tubing - ASTM D 3295

Standard of the tubing. It is intended for electrical, mechanical, chemical and medical applications manufactured from extrusion resins made from PTFE resins.

Molding and Machining Tolerances for PTFE resin parts - ASTM 3297

This standard specification defines tolerances applicable to parts molded and free sintered from PTFE resins and to machine parts produced from basic shapes of compression molded or ram extruded resins. The thermal expansion of PTFE parts between 65 degree of F and 69 degree of F is non uniform due to a critical transition zone characteristic of PTF**E.**

Importance of Filler Materials in PTFEs:

Shortcomings of PTFE, if used alone for designing of a recip. gas compressor parts:

Most of the reciprocating gas compressors are installed in the oil fields and related industries. So, they have to work in the dry conditions for most of the times of their life. PTFE alone becomes brittle in this kind of the dry friction atmosphere. So, some fibers with PTFE materials are to be incorporated to have strength with the wear resistance at the elevated temperatures and the pressures. Self lubrication properties also become helpful for this issue. That's why some of the companies had developed the filled PTFE materials for their particular applications. We will see in the next sessions various fillers used in the PTFE and their properties also.

Filled PTFEs:

Various fillers can be added with the PTFE as a base/matrix resin to enhance certain properties, e.g. glass fiber, glass bead, carbon, graphite, molybdenum disulfide, bronze, etc. PTFE does not melt; it cannot be molded into complex shapes, but must be machined. PTFE is easily machined using standard mechanical woodworking and stamping equipment and tooling. Most shapes are sold slightly oversized for easy trimming and machining to exact sizes and shapes with dimensional accuracy and to achieve the close tolerances

Contribution of the filler materials' functions and properties to the filled PTFEs:

1) Glass fibers:

If PTFE is reinforced with glass fibers, the percentage will be varying between 8 and 38%. The added glass fiber improves the wear properties to a minor degree, also the deformation strength under load while leaving substantially unchanged the electrical and chemical characteristics. The coefficient of friction is slightly increased for the resulting material and for this reason; graphite is sometimes added to compensate this side effect. The diameter should be kept small for achievement of flexibility along with the long whiskers shape.

2) Carbon fibers:

Carbon is added to the PTFE in a percentage by weight between 12 and 32%, along with small percentage of graphite. Also, the carbon tends to improve to a considerable degree of wear resistance and the deformation strength, while leaving practically unchanged the chemical resistance. It is substantially modifying the electrical properties.

3) Bronze fibers:

Bronze is added as a feeler in percentages of weight between 42 and 58%. Bronze filled PTFE has the best wear properties, remarkable deformation strengths and good thermal conductivity but poor electrical characteristics and chemical resistance.

4) Graphite fibers:

The percentages used for these fibers vary between 7 and 14%. Graphite lowers the coefficient of friction and is, therefore, often added to other types of filled PTFE for improving this property. It improves the deformation under load, strength and, to a minor degree, the wear resistance properties. The diameter should be kept small for achievement of flexibility along with the long whiskers shape.

5) Other fillers fibers:

Molybdenum with sulfide coating, though decreasing the coefficient of friction, is sometimes preferred over the graphite. Some metal powders (stainless steel, nickel, titanium), in consideration of their particular resistance to the chemical agents are sometimes used as fillers for PTFE, even though their wear resistance, with respect to bronze, are inferior. The metal oxides are added to other fillers to give the better wear resistance properties.

Mechanical properties after fillers added to the PTFEs:

1) Wearing resistance properties:

The contact between two sliding surfaces, because of the friction is generated in the contact zone, which results in a

certain wear. The magnitude of this wear depends on the load, speed and time of sliding contact. Theoretically, this phenomenon can be described by the equation given as per below:

R = PVTK

Where, expressed in the measuring units of table:

R = wear in mm

P = specific load in N/mm2

V = sliding speed in m/sec

T = time in hrs

K = wear factor in (mm3 .sec)/(N.m.hrs)

The value of the factor PV after which the coefficient of wear loses its linear behavior, assuming remarkable values with the system passing from weak to strong wear condition, is known as PV limit. This PV limit and the wear factor are, therefore, characteristic parameters of choice for each material. In practice, however, it can be easily perceived, the wear factor and the PV limit of the same filled material can vary also with the nature, the hardness and the surface finish of the other contact partner component with the presence, or not, of cooling and lubricating fluids.

2) Deformation under load and compressive strength:

PTFEs like most other plastic materials - has no elastic zone where the ratio of load/deformation has a constant value. This ratio of load/deformation depends upon the time of application of the load and the ensuing deformations; this phenomenon is known as creep and at the removal of the load, there is only a partial return of the deformation to the original state of elastic recovery, so that we are always in the presence of a permanent deformation. Creep, obviously not being a linear function of time, results after just over 24 hrs in deformations which in most cases are not taken into consideration. With increasing temperature, there is a falling off of the deformation under load properties and consequently of the compressive strength which is already at 100°C equal to half of that at 296.16 K and at 473.16 K about 0.1. In any case, PTFE, and in particular filled PTFE, is one of the plastic materials retaining the properties at high temperatures, optimum deformation properties under load. To conclude, the elastic recovery in about half percentage of the deformations under load and the permanent deformations are equal to about half percentage of the deformations under load. This applies both to filled and unfilled PTFE. The properties of the first are superior then second. In fact, the deformation under load of the more common types of filled PTFE are about quarter of those of the unfilled ones, while the compressive strength is about twice as high.

Thermal properties after fillers added to the PTFEs:

The COTE of filled PTFE is in general inferior to that of unfilled PTFE and always greater in the direction of the molding. The thermal conductivity is superior to that of unfilled PTFE, particularly when using fillers having their higher thermal conductivity. Filled PTFE therefore have better thermal properties than the unfilled.

Electrical properties after fillers added to the PTFEs:

These properties depend to a large degree upon the nature of the filler fibers. Only PTFE filled with glass fiber possesses good dielectric properties, even though different from those of unfilled PTFE. For example, the volume and surface resistivity, the dielectric constant and the dissipation factor may vary largely with the variation of the humidity of an atmosphere and operating/electrical frequency. These properties are very much concerned to this field because most of the compression systems are working in the fire hazardous atmosphere and most of them are having electrical motor as a driver and starters for the motors are also closely connected to the system. A single spark can cause the entire plant shutdown and causes the unbelievable losses in terms of human lives and capital losses.

Manufacturing and Testing Methods for Filled PTFE Wear Parts like Piston Rings, Rider Rings and Packing Rings for Reciprocating Gas Compressors. Most of the parts are made up from the raw materials like filled PTFE balls available in the market. To have additional materials deposition on to the surface the various etchants are used and then high pressure and temperatures are applied to the filler materials to have proper adhesive bonding between the fillers and the PTFEs. However, the main methods comprises of the following methods to produce shapes like tubes, pipes etc.

Automatic pre – forming shapes from balls by compression molding

Multi cavity protrusion and extrusion processes

Modified fusion welding for filled PTFE

Transfer molding

Injection molding

For producing final shape with required dimensions of the rings, machining is to be done. The material has good machining characteristics. Today the CNCs can efficiently machine these materials and they found economical too. However, the alternative methods of production like forming; pressing etc. in the metals industries cannot be applied to these materials. Also, the field of the non destructive testing is preferred for testing of the products in the industry over the destructive one. Most of the testing comprises of x-ray, dye penetrate test, surface projection test, ultrasonic test etc. For destructive testing – the tensile as well as compressive loading test along with the wear measuring test is carried out.

III- ECONOMICAL ECLAIRS CRITERIA

There are several criteria to have a selection of the newly developed material for the use and production purpose over the existing material. We have seen the success of filled PTFEs over the metallic materials for the wear parts as per technical point of view. That was the first criteria. The other criteria are the cost effectiveness and consistent commercially supplements over the long span of time with easiness and short time period installations.

Criteria # 1

Cost Effectiveness & Economic Batch Quantity / Mass Production:

Most of the metals used in making of the wear parts are bronze, aluminum, k-monel, inconel, steel, cast iron and various alloys. They are available in the market with very high price per pound as compared to the filled PTFEs and that too in the form of raw

materials. If the finished components are to be considered, the filled PTFEs are much cheaper than the other metallic materials. For example the bar of 2.2 pound pure bronze is for sell for $ 5.80 approximately, while the finished ring of the glass filled PTFE is available in $ 3.86 in the market.
Industry had already started the mass production for some of the reciprocating gas compressor manufacturing companies, having the patented name of the high performance wear parts. They have incorporated couple of analysis to have a successful production and higher market demand, amongst them one is the economic batch quantity analysis for the mass production and inventory control purpose, presented here. According to this theory the economic batch quantity can be given by the formula given below:

$$\frac{d(TC)}{dQ} = \frac{C_C(p-r)}{2p} - \frac{C_p A}{Q^2} = 0$$

Where, Q = economic batch quantity
p = daily production rate
Cp = cost of purchase from raw material balls
Cc = carrying cost
r = usage rate

Thus, as per the costing analysis this material is found cheaper than the metallic one.

Criteria # 2

Commercially Consistent Supplements Over the Long Span of Time with Easiness and Short Time Period Installations:

We have seen that the compression systems have to work in the fire HAZOPE zones and are mostly employed in petrochemical industry, where breakdown serviceability of these parts become even more critical. So the consistency in the availability of the same material with same properties for over long span of time is required with the easiness and short time period installations without failure during assembly. Achieved flexibility due to composites and synthetically made polymers satisfy both of these requirements over the metallic parts. The rate of breaking of parts during on site installations or on site replacements is found higher in the metallic rings than the filled PTFEs. Also, earth crust has negligible amount of the metals in, from them the fraction of the required metals for making of these parts is very less as compared to we can synthesize the polymers and use the other fillers for making these filled PTFEs.

IV-CASE STUDY

Here one of the piston' rider ring design case is proved for the betterment use of the filled PTFEs over the conventional metallic ring. The ring sizing program2 is used to determine the size of the ring and using AutoCAD® the mass for the part is calculated and the result found out was the glass filled PTFE is much lighter than the bronze one. However, other criteria like cost analysis, detailed design analysis, etc. are excluded.

Here, the piston sizing calculation program made by me, while working with the Dresser-Rand India Pvt. Ltd. is shown in figure[3].

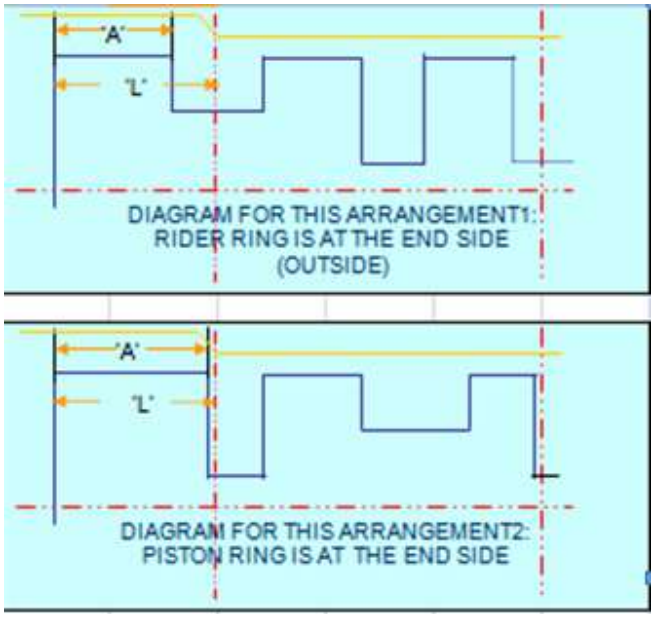
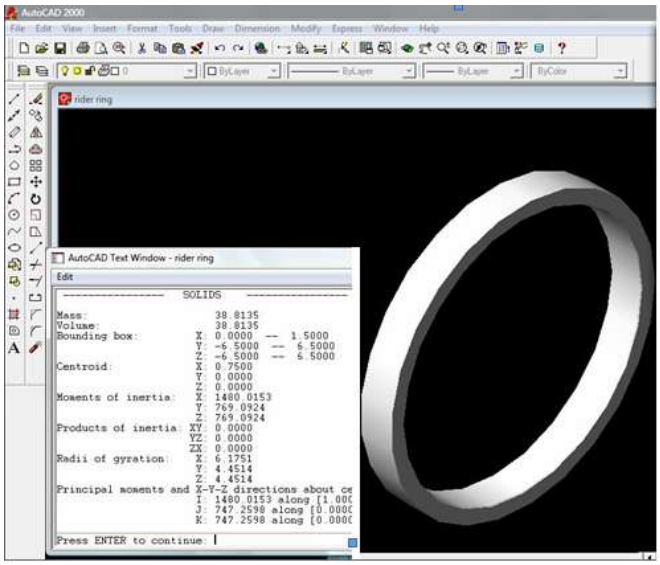


FIG 3: piston sizing program and piston ring rider ring arrangement

Here, as shown in figure[4] the ring size came up to 0.625" X 1.500", which has developed in the AutoCAD® and the volume of the ring was found out like 38.8135 inch^3. If this is multiplied by the density of the bronze (0.3 lbs/ inch^3) than the mass comes 11.64 lbs, whereas, if it is multiplied by the density of the 29% filled PTFE (0.079 lbs/inch^3) than the mass comes 3.07 lbs, which is 3.8 times lighter than the metallic one for only one ring. The total reciprocating mass of wear parts make huge difference and reducing the unbalanced forces and vibrations, increasing the product's life, system's life etc. all advantages are incorporated with it.



V. CONCLUSION

The new composites and compounds of filled PTFEs are proved the best replacement substitutes for metallic wear parts like piston rings, rider rings and packing rings in the reciprocating gas compression systems in all aspects like,

- Weight Reduction of the Reciprocating Parts
- Vibrations Reduction of the Compressor
- Better Balancing of the Compressor
- Cheaper and Better Machining of the Wear Parts
- Economical Maintenance of the Compressor and Wear Parts
- Improved Lifecycle of the Wear Parts
- Less Production Time of the Wear Parts
- Easy and Faster Transportation of the Wear Parts
- Easy Assembly of the Compressor Parts and Wear Parts
- Suitability with Lube, Non Lube and NACE Applications

The industries are implementing this change rapidly to overcome the disadvantages incorporated with the older metallic materials for the wear parts and to gain the advantages incorporated with newer filled PTFEs, to cope up with the time line demands of the rapidly developing industries.

ACKNOWLEDGEMENT

Author Prof. Kedar Bhojak gives his sincere vote of thanks to Dolores Foster, Prof. Gregory Ojard, Prof. Durgesh Jani and Prof. Bhoomi Trivedi for their technical and moral support.

REFERENCES

K. Chawla, *Composite Materials: Science and Engineering*, 2nd Edition, Springer Science Business Media, Inc., 1998.

4G Wireless Technology

\Ms. Ketaki Sutariya
Asst. Prof.,
Electronics & Communication Dept.,
Sabar Institute of Tech. for Girls, Tajpur
E-mail: ktk_21_06@yahoo.co.in
Ms. Kruti Lavingia
Lecturer,
Information Technology Dept.,
Sabar Institute of Tech. for Girls, Tajpur
E-mail: krutilavingia@gmail.com

***Abstract*— Today, mobile communications play a central role in the voice/data network arena. With the deployment of mass scale 3G just around the corner, new directions are already being researched. In this paper we address about the 4G mobile communications.**

The advent of 4G wireless systems has created many research opportunities. The expectations from 4G are high in terms of data rates, spectral efficiency, mobility and integration. Orthogonal Frequency Division Multiplexing (OFDM) is proving to be a possible multiple access technology to be used in 4G. But OFDM comes with its own challenges like high Peak to Average Ratio, linearity concerns and phase noise. This paper proposes a solution to reduce Peak to Average Ratio by clipping method. MATLAB as used to generate the OFDM signal to prove that clipping does reduce Peak to Average Ratio.
This paper presents an overall vision of the 4G features, framework, and integration of mobile communication. The features of 4G systems might be summarized with one word—integration. The 4G systems are about seamlessly integrating terminals, networks, and applications to satisfy increasing user demands. The continuous expansion of mobile communication and wireless networks shows evidence of exceptional growth in the areas of mobile subscriber, wireless network access, mobile services, and applications.

***Index terms*—4G, OFDM (Orthogonal Frequency Division Multiplexing), Spectral Efficiency**

I. INTRODUCTION

The first operational cellular communication system was deployed in the Norway in 1981 and was followed by similar systems in the US and UK. These first generation systems provided voice transmissions by using frequencies around 900 MHz and analogue modulation.
The second generation (2G) of the wireless mobile network was based on low-band digital data signaling. The most popular 2G wireless technology is known as Global Systems for Mobile Communications (GSM). The first GSM systems used a 25MHz frequency spectrum in the 900MHz band.

Planning for 3G started in the 1980s. Initial plans focused on multimedia applications such as videoconferencing for mobile phones. When it became clear that the real killer application was the Internet, 3G thinking had to evolve. As personal wireless handsets become more common than fixed telephones, it is clear that personal wireless Internet access will follow and users will want broadband Internet access wherever they go.

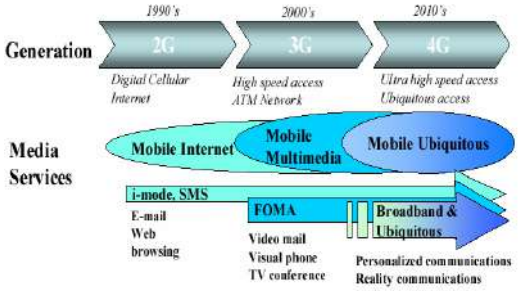


Fig. 1 Evolution of Mobile Communication Systems

The objective of the 3G was to develop a new protocol and new technologies to further enhance the\mobile experience. In contrast, the new 4G framework to be established will try to accomplish new levels of user experience and multi-service capacity by also integrating all the mobile technologies that exist (e.g. GSM - Global System for Mobile Communications, GPRS - General Packet Radio Service, IMT-2000 - International Mobile Communications, Wi-Fi - Wireless Fidelity, Bluetooth).In spite of different approaches, each resulting from different visions of the future platform currently under investigation, the main objectives of 4G networks can be stated in the following properties:
• Ubiquity;
• Multi-service platform;
• Low bit cost

To achieve the proposed goals, a very flexible network that aggregates various radio access technologies, must be created. This network must provide high bandwidth, from 50-100 Mbps for high mobility users, to 1Gbps for low mobility users, technologies that permit fast handoffs, an efficient delivery.

Unlike previous standards such as 3G (third generation), 4G is based entirely on packet switched networks. In addition, all 4G networks will be digital and will provide higher bandwidths of up to 100 Mbps. 4G is actually a collection of previous standards as oppose to an entirely new standards. Standards such as 3G and Bluetooth will be incorporated in to the 4G standards.

II. ACCESS TECHNOLOGY

The move towards what is being coined fourth generation (4G) wireless is complicated by the fact that a single 3G standard upon which to build does not exist. However, most industry experts agree that the future of wireless is one in which voice, video, multimedia and broadband data services traveling across multiple wireless air interfaces are meshed into one seamless network. 4G Wireless networks will be characterized by the following:

1. Seamless network of multiple air interfaces and protocols
2. Improved spectral efficiency
3. IP Based
4. Higher data bandwidths (<100Mbps)

Several technologies based on multi -carrier modulation (MCM) have come to the forefront in order to achieve the above characteristics. We shall examine one such technology known as Orthogonal Frequency Division Multiplexing (OFDM). But the fact that OFDM is just one of many competing technologies highlights the need for reconfigurable and flexible software defined radio systems as a development platform. Here we will present architecture for a software development platform by first looking at the challenges of designing an OFDM -based system.
OFDM, orthogonal frequency division multiplexing, is a telecommunications technology that is the foundation of most next-generation, or 4G wireless Internet services for high speed data transmissions at comparable rates to fixed DSL services. OFDM transmits multiple signals simultaneously over a single wireless system.

This successful implementation of the OFDM waveform is the first step in Military SDR Technologies plan to implement the complete IEEE 802.16 family of wireless data applications, also known as WiMax, a higher performing version of WiFi. Emerging as the predominate technology for providing cost-effective, high-speed wireless Internet solutions, WiMax capabilities compare to traditional land based DSL services in terms of speed and cost, with the added capability of wireless mobility. WiMax-based services will be capable of delivering high quality voice, video and multimedia content over an Internet-based connection.

OFDM (Orthogonal Frequency Division Multiplexing)

OFDM is a communications technique that divides a communications channel into a number of equally spaced frequency tones (bands). OFDM is a form of multi-carrier modulation (MCM) where a sub-carrier within each frequency band is modulated to carry a portion of the user information. A communications data stream is effectively split into N parallel low bandwidth modulated data streams (Figure 1). Each sub-carrier overlaps, but they are all orthogonal to each other, such that they do not interfere with one another.

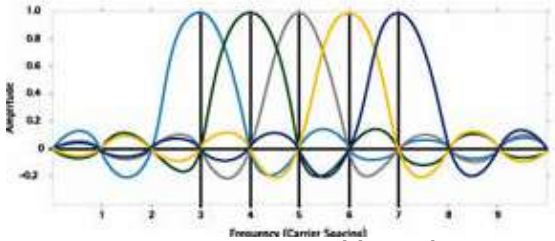


Fig. 2 OFDM Modulation Scheme

Each of the sub -carriers has a low symbol rate. But the combination of sub-carriers carrying information in parallel allows for high data rates. The other advantage of a low symbol rate is that inter-symbol interference (ISI) can be reduced dramatically since the symbol time represents a very small proportion of the typical multipath delay. The transmitter stage of an OFDM transceiver takes data from an IP network, converts, and encodes it into a serial stream before modulation. The OFDM signal is generated using an Inverse Fast Fourier Transform (IFFT) into an IF analog signal which is then sent to the RF transceiver. The receiver stage of the transceiver simply reverses the process.

OFDM provides a particularly robust air interface that is resistant to the effects of multi-path delays while maintaining spectral efficiency. However, designers of OFDM and other MCM-based systems have been forced to deal with a host of challenges.

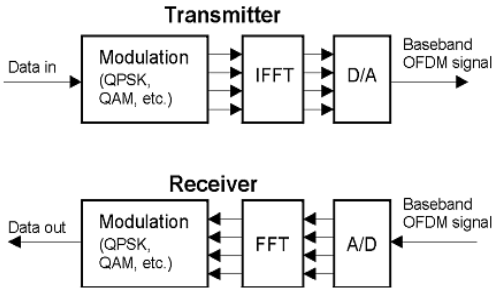


Fig. 3 Basic OFDM Transmitter and Receiver

IV. 4G NETWORK ARCHITECTURE

The number of access networks in public, private business and home areas is increasing. To save both users and services from having to deal with many different access technologies, authentication strategies, and network peculiarities, these details should be hidden from them.

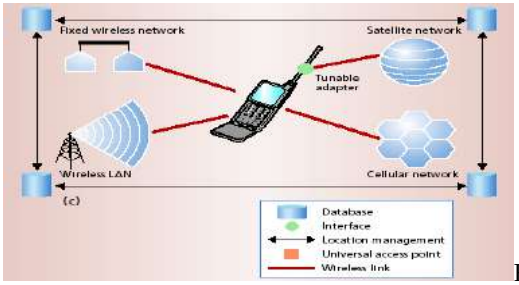


Fig. 4 Possible 4G Network Architecture

On the other hand, the characteristics of the particular access network being used do matter. End users for example want to be able to use a specific network based on certain selection criteria like speed and cost. Service providers will need access to specific network characteristics to be able to enhance their services, and may have certain requirements such as minimum delay guarantees. For the end user this means that in most circumstances he will not be aware of his services using different access networks simultaneously. This involves concepts like seamless roaming and the Virtual Home Environment (VHE). Another aspect is that the user needs to control the usage of the available networks, especially when this usage comes with a price. This involves a (potentially complex) decision making process which may be guided by policy management tools, with support from both end user devices and the networks. Another kind of integration is that of public, private and home networks. This has not received much attention yet, at least not to the extent as expected to be needed for next generation networks.

Fig. 5 Business model of 4G Network

The current model assumes there are relationships between the end user, the service provider, and the network operator. In the 4G worlds, the number of access networks and therefore the number of network operators will increase rapidly. Especially the construction of many so-called “hotspots”, urban areas with Wireless LAN coverage, will cause this. Both the end user and the service provider do not want to be bothered with the peculiarities of these networks, i.e. end users still want to use the services offered by service providers irrespective of the network they are connected to. This asks for a service centric approach where the end user, the service provider and network operator are loosely coupled through a Service Support role. Two main responsibilities of Service Support are service aggregation and network integration Service aggregation (SA) enables ubiquitous service provisioning to end users, whereas network integration (NI) ensures that this can be realized regardless of the network the end user is connected to. The end user benefits from being able to access his/her subscribed services anywhere and anytime.

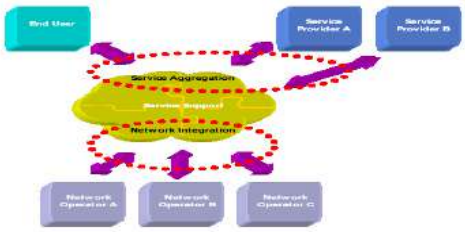


Fig. 6 Service Centric Approach

IV. ADVANTAGES OF 4G

4G, the latest generation in wireless technology, offers you connection speeds like you have never experienced before and a quality, reliable signal that will never let you down. Fourth generation connections are nearly four times faster than their predecessors; to put this into perspective, you can easily download an entire feature-length movie in mere minutes, instead of hours. Never again do you have to let slow download speeds or the glacially-paced loading of data-rich websites drag you down. With fourth generation connections, you can maneuver the internet as it was meant to be experienced, at super-fast speeds and with the agility to jump from one program or page to the next without skipping a beat. Even watch live streaming television shows with the picture clarity and quality you can expect from a satellite provider, something 3G connections could never deliver.

In addition to being clearly superior in terms of connection speeds, fourth generation provides a much more reliable signal than any previous service has ever been able to offer. You know how frustrating it can be when you are trying to finish a big project for school or work, and your internet connection is on the fritz; it can really cause a lot of unneeded stress or panic. With fourth generation, you can rest easy knowing that your internet connection will be there for you whenever you need it. In addition, it is also much safer than connecting over a traditional Wi-Fi signal. The connection is much more secure and you don't have to worry about hackers getting into the signal and gaining access to your personal files and information. Fourth generation connections not only provide a strong, reliable signal, they are incredibly secure as well.

Another big advantage of 4G wireless is the huge range of coverage areas it can service. Instead of providing access to the internet to a household, apartment building or even a city block, fourth generation signals can reach across entire cities, providing you with a secure and fast internet connection wherever your busy day may take you. This newest technology will change the way you use the internet for the better and allow you to unlock its full potential. With so many clear benefits, what are you waiting for? Upgrade today to the fourth generation and see what you've been missing.

V. CONCLUSION

It is always dangerous to predict too far ahead in a fast-moving field such as mobile communications. Almost by definition the eventual 2010 scene will not match exactly that depicted in the 4G vision described herein. However, the key elements—fully converged services, ubiquitous mobile access, diverse user devices, autonomous networks and software dependency—will persist. The 4G Vision is a living document which intends to update and amend as time and knowledge progress. It will act as the umbrella vision to a large research programme and place in context the detailed research work that will take place in the various areas. In this respect it will help to continuously steer the research as it progresses and, therefore, to make it more relevant and beneficial.

REFERENCES

[1] TUTFLEBEE, W. H. W.: 'Mobile VCE: the convergence of industry and academia', Electron. Commun. Eng. J., December 2000, 12, (6), pp.245—248

[2] IRVINE, J., et al.: 'Mobile VCE scenarios'. A document produced as part of the Software Based Systems work area within the Mobile VCE Core 2 research programme, September 2000. See http://www.mobilevce.com

[3] 'Wireless Strategic Initiative'. An EU project. See http://www.ist-wsi.org

[4] Foresight ITEC Group: 'Visions'. See http://www.foresight. gov.uk

[5] Jun-seok Hwang, Roy R. Consulta & Hyun -young Yoon, '4G Mobile Networks – Technology beyond 2.5G and 3G'

Comparison of Performance Ratings of EDFA and EYCDFA for 1535 nm to 1580 nm Signal Wavelength Span

Ronak R. Vashi, Pravin R. Prajapati

***Abstract*— This paper presents for broad-band wavelength division multiplexing (WDM) amplification, Erbium Ytterbium Co Doped Fiber Amplifier (EYCDFA) is a better option than Erbium Doped Fiber Amplifier (EDFA) and for single channel transmission EDFA is better one. Relationship between various parameters like gain, noise-figure and output optical power for EDFA and EYCDFA is analyzed. Source wavelength of 1510 nm to 1580 nm, fiber length of 14 meter, pump wavelength of 980 nm and -10 dBm input source power are considered. Due to the small source power, Rayleigh back scattering is neglected. For simulation of optical link, licence version of OPTSIM 5.2 is used.**

***Index Terms*— EDFA, EYCDFA, ytterbium, erbium, wavelength division multiplexing**

I. Introduction

Many recent different communication links have employed optical fiber in their structure, mostly due to the advantageous that the fiber can deliver to the system, such as wide bandwidth, low loss, low distortion and electromagnetic isolation [3]. It was only after the development of optical amplifiers using rare-earth doped optical fibers that systems with high capacity of transmission could become commercially available.

Optical fiber amplifier with broad-band amplification has become more important with the continuously increasing number of wavelengths channels in optical network systems. One critical issue for wavelength-division multiplexing (WDM) networks is the need for gain control to the limit the effect of multichannel cross saturation [1-3], i.e., the dependence of the gain in any one channel on the total input power in all channels.

Erbium Doped Fiber Amplifier (EDFA) has major disadvantage is different gain at different wavelength, so we cannot use as WDM application. Several methods have been proposed for eliminating the unwanted dependence of gain on the number of channels in multiwavelength EDFA's [4].

Form that Erbium Ytterbium Co Doped Fiber Amplifier (EYCDFA) having flattened gain curve for specific wavelength region, so it becomes more popular although it provides higher Noise figure than EDFA. EYCDFA has demonstrated high and broad gain for particular wavelengths, from conventional EDFA.

In Yb^{3+} -co doped EDFA's, the Yb^{3+} ions absorb most of the pump power and cross correlation between adjacent of Yb^{3+} and Er^{3+} ions ensures the energy transfer. The introduction of Yb^{3+} allows for higher gain for given pump level [6].

For a practical application of the model, the effect of the Yb^{3+} - Er^{3+} ratio and Er^{3+} concentration on the fiber gain and the optimum fiber length. The selection of the optimum fiber length for a particular application is a key design effort [9]. The optimum fiber length is defined as the fiber length for which the local net gain remains positive, i.e., the pump power still provides enough gain to overcome the losses.

MATHEMATICAL MODEL OF EDFA AND EYCDFA

The simplest treatment of the EDFA and EYCDFA starts out by considering a pure three-level for EDFA and five-level for EYCDFA atomic system. Most of the important characteristics of the amplifier can be obtained by this mathematical model.

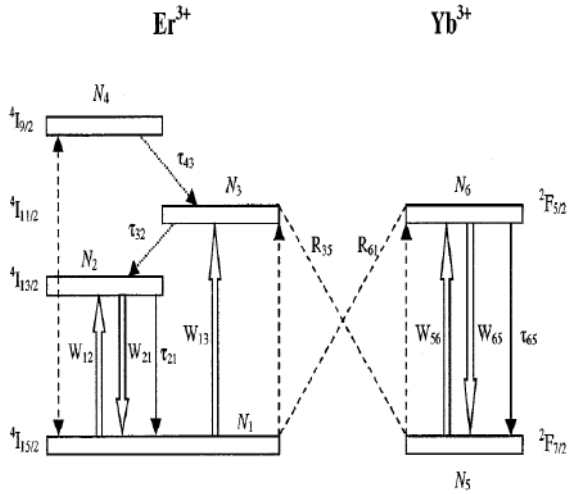


Fig. 1 Energy level transitions for Er^{3+} and Yb^{3+} systems.

A. EDFA

Manuscript received on March 28, 2011. This work is supported by the A.D. Patel Institute of Technology, NewV.V.Nagar, Gujarat.

Ronak Vashi is the P.G. student of G.H.Patel College of Engineering and Technology, V.V.Nagar, Gujarat, India. (Phone: +919426724335;E-mail: vashi_ronak@yahoo.co.in)

Prof. Pravin R. Prajapati is Associate Professor, Electronics and Communication Department of A.D.Patel Institute of Technology. (E-mail: p_r_prajapati@yahoo.com)

The rate equation for the population changes are written as [6].The notation and basic equation for the model similar to the equations given in [2]. Due to experimental findings the rate equations are modified.

As shown in Fig.1 with a ground state denoted by 1, state 2 is the upper level of the amplifying transition and state 1 is lower level and energy is pumped in state labeled 3. The population's levels are labeled N_1, N_2, and N_3. The signal absorption and signal emission to the 1 to 3 transition is denoted by W_{13} and the incident light flux at the frequency corresponding to the 1 to 2 transitions and cross-correlation coefficient is denoted by R_{12}. In case of the Er^{3+} ions, $^4I_{13/2}$ (level 2) to $^4I_{15/2}$ (level 1) transitions occurred [3].

$$\frac{dN_3}{dt} = -\frac{N_3}{\tau_{21}} + W_{13}N_1N_3 \tag{1}$$

$$\frac{dN_2}{dt} = -\frac{N_2}{\tau_{21}} + \frac{N_3}{\tau_{32}} - N_2N_1R_{21} \tag{2}$$

$$\frac{dN_1}{dt} = \frac{N_2}{\tau_{21}} - W_{13}N_1N_3 + N_2N_1R_{21} \tag{3}$$

In a steady-state situation, the time derivatives will be zero,

$$\frac{dN_1}{dt} = \frac{dN_2}{dt} = \frac{dN_3}{dt} = 0 \tag{4}$$

And the total population N is given by [1]

$$N_{Er} = N_1 + N_2 + N_3 \tag{5}$$

B. EYCDFA

Amplified spontaneous emission (ASE) due to $^2F_{5/2}$ - $^2F_{7/2}$ transitions at 980 nm pump wavelength, as well as in the vicinity of 980 nm wavelength due to $^4I_{11/2}$- $^4I_{15/2}$ transitions, is considered in the model. In addition, signal ASE due to 1510 nm to 1580 nm wavelength span is considered [6].

$$\frac{\partial N_3}{\partial t} = -\frac{N_3}{\tau_{32}} + W_{13}N_1 - R_{35}N_3N_5 + R_{61}N_1N_6 + CN_2^3 \tag{6}$$

$$\frac{\partial N_6}{\partial t} = -\frac{N_6}{\tau_{65}} + W_{56}N_5 - W_{65}N_6 + R_{35}N_3N_5 - R_{61}N_1N_6 \tag{7}$$

$$N_{Yb} = N_5 + N_6 \tag{8}$$

Where the W_{ij} terms represent the stimulated rates between level i and j of signal absorption, signal emission, Er^{3+} and Yb^{3+} pump absorption and pump emission rates, R_{ij} is the cross section coefficients and C upconversion coefficient.

For the input and output signal powers [5],

$$P_{p,s}^{out} = P_{p,s}^{in} e^{-\alpha_{p,s}L} e^{\left(P^{in} - P^{out}\right)/P_{p,s}^{sat}} \tag{9}$$

$$P_s(sat) \approx P_p^{out}(0)\left[1 - e^{\left(P_s^{sat}/P_p^{sat}\right)}\right] \tag{10}$$

P^{in}, P^{out} are the total input and output powers, P^{sat} is saturation power of EDFA and $P_p^{out}(0)$.

Where, $\alpha_{p,s}$ is the attenuation constant and $P_{p,s}^{sat}$ saturation power for pump signal can be obtained [2] from eq. (11) and (12).

$$\alpha_{p,s} = \rho\Gamma_{p,s}\sigma_{p,s}^{a} \tag{11}$$

$$P_{p,s}^{sat} = \frac{A}{\left(\sigma_{p,s}^{e} + \sigma_{p,s}^{a}\right)\tau\Gamma_{p,s}} \tag{12}$$

The analytical model has been shown to be quite accurate for predicting gain values in EDFA for gain less than 20 dB, where saturation of the amplifier by ASE is not significant [3].

$$G(dB) = 10\log_{10}\left(I_s(z=L)/I_s(z=0)\right) \tag{13}$$

The input and output powers of an EDFA can be expressed in terms of the principle of energy conversion [2]. In addition, pumping at 980 nm wavelength is preferred, since it is produces less noise and achieves larger population inversion than pumping at 1480 nm wavelength.

The noise figure (NF) is the terms of the ASE power exiting the fiber in bandwidth Δν [3],

$$NF\,(dB) = 10\log_{10}\left(\frac{P_{ASE}}{h\nu\Delta\upsilon G} + \frac{1}{G}\right) \tag{14}$$

Following parameters of fiber are used for simulation of optical link.

TABLE 1
PARAMETER USED IN THE MODEL

Parameter	Symbol	Value	Unit
Er^{3+} ion concentrations	N_{Er}	1.35×10^{25}	m^{-3}
Yb^{3+} ion concentrations	N_{Yb}	8.50×10^{25}	m^{-3}
Spontaneous Emission Lifetime form level 2 to 1	τ_{21}	11.0	ms
Spontaneous Emission Lifetime form level 3 to2	τ_{32}	6.7	µs
Spontaneous Emission Lifetime form level 6 to 5	τ_{65}	0.35	ms
Cross-correlation coefficients	R_{35} R_{61}	3.0×10^{-22}	m^3/s

II. DETAILS OF SIMULATION SET UP

In this simulation, various parameters of both amplifiers like gain, noise figure and output optical power is compared. 14 metre EDFA with an input signal power of –10 dBm, and pump wavelength 980 nm is considered.

For simulation current topology where co-pumping propagates with the signal is considered, source wavelength vary from 1510 nm to 1580 nm in step of 10 mm.

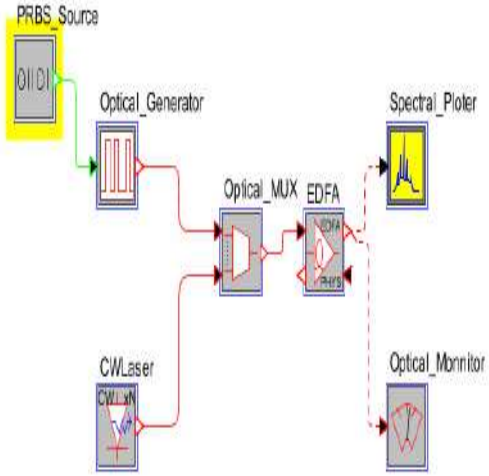


Fig. 2 Block diagram of the basic EDFA configuration.

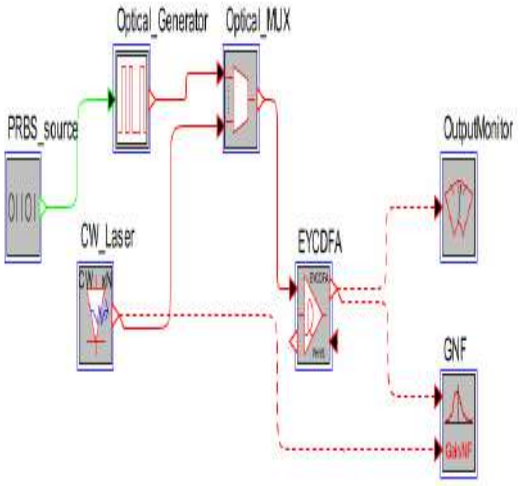


Fig. 3 Block diagram of the basic EYCDFA configuration.

III. RESULTS AND DISCUSSION

A. *Gain*

Fig. 4 shows the signal gain profile for the (a) co-propagating EDFA configuration and Fig.5 shows EYCDFA configuration when signal wavelength varies from 1510 nm to 1580 nm. EDFA gives higher gain of 40 dB at 1535 nm. As wavelength increase from 1530nm to 1580nm, gain of EDFA rapidly decreases. EYCDFA gives almost constant gain of 30 dB for 1550nm to 1580 nm input signal.

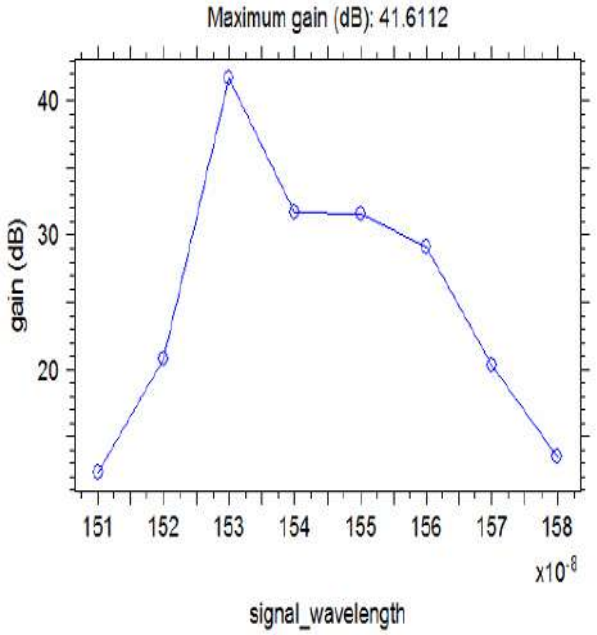


Fig.4 Gain(dB) v/s source wavelength(m) of EDFA

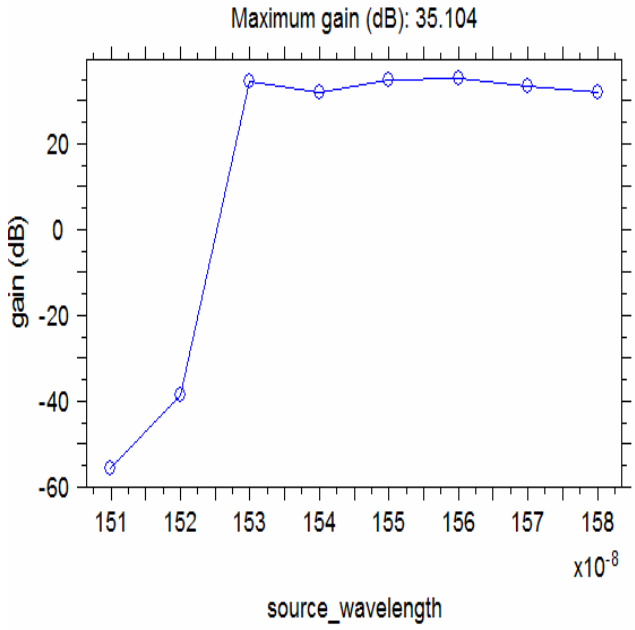


Fig. 5 Gain v/s source wavelength of EYCDFA.

B. *Noise Figure*

As shown in Fig. 6 and Fig.7, for a 14 meter fiber of EDFA and EYCDFA noise figure improves monotonically from 1510 nm towards the longer signal wavelengths. Since the absorption is stronger at 1530 nm than at 1550 nm, these will improvement in noise figure. EDFA gives better noise figure than EYCDFA for 1510nm to 1580 nm source wavelength region.

The noise figure shows only a weak dependence on variation in the Er^{3+} concentration or Yb^{3+} and Er^{3+} ratio. For EYCDFA average noise figure was about 6 dB, with a maximum variation of 2.4 dB.

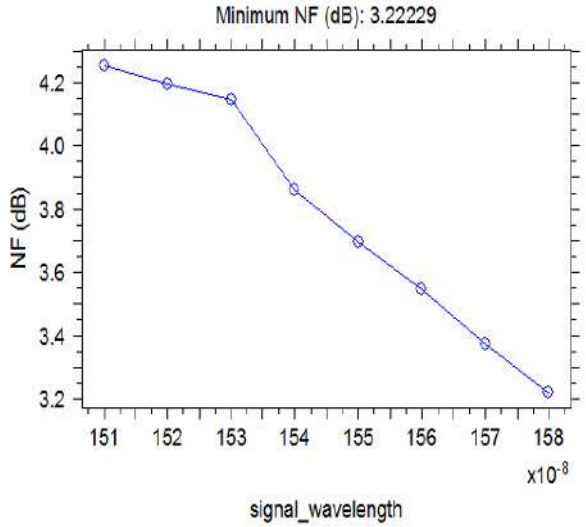


Fig. 6 Noise Figure(dB) v/s Source wavelength (m) of EDFA.

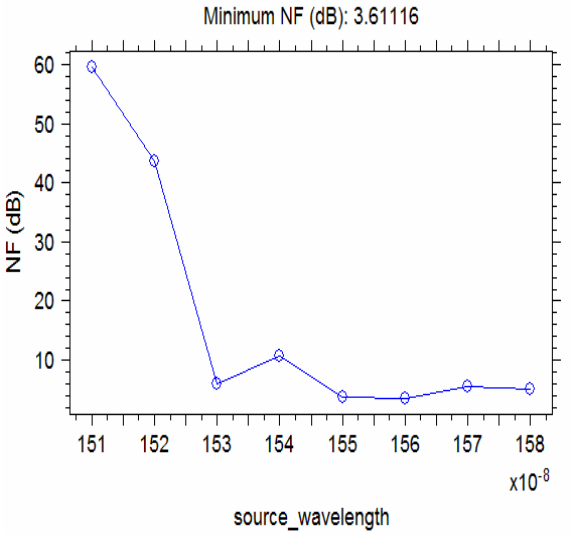


Fig. 7 Noise Figure(dB) v/s source wavelength (m) of EYCDFA.

C. *Output Optical Power*

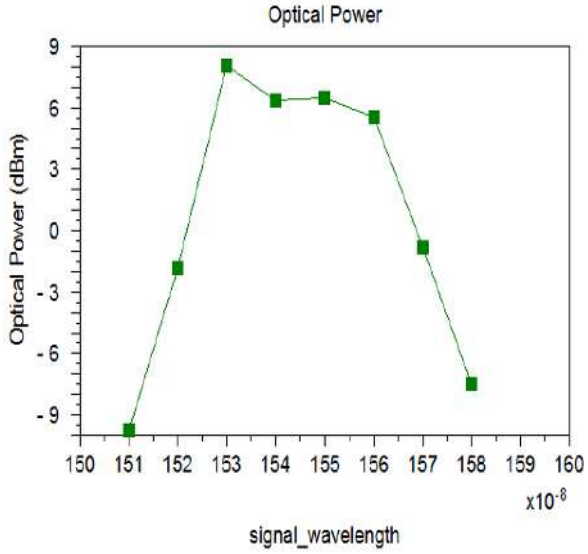


Fig.8 Optical Power(dBm) v/s Source wavelength(m) of EDFA

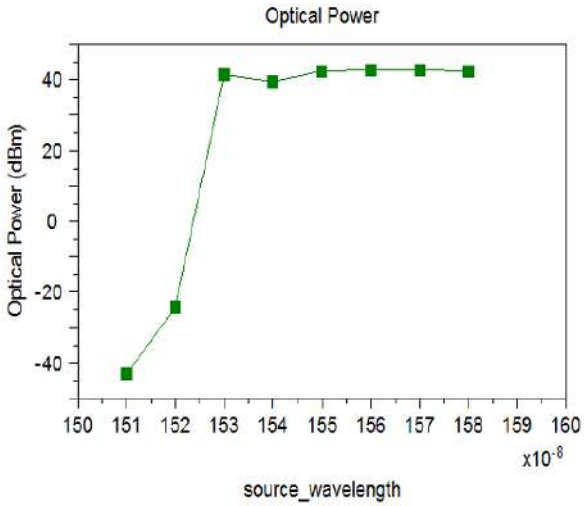


Fig.6 (b)

Fig. 9 Optical Power(dBm) v/s source wavelength (m) of EYCDFA

Higher output power can be obtained from ytterbium-doped fiber as per Fig. 9. EYCDFA gives approximate constant output optical power of 40 dBm for 1550 nm to 1580 nm input signal wavelength span. In fact, this behavior accepted because in the saturated regime, it is possible to convert quite a larger fraction of pump energy into signal energy, and it is much more than that of EDFA.

V. Conclusion

At 1535 nm EDFA gives maximum gain as well as lower noise figure. So, single channel single mode fiber EDFA is best suitable as pre-amplifier means before receiver, where lower noise figure and higher gain is required.

For wavelength division multiplexing (WDM) application, where there is transmission of more than one wavelength, constant gain of amplifier and constant output optical power is required over large span of wavelength. EYCDFA gives constant curve for gain and noise figure for 1550 nm to 1580 nm input wavelength.

Acknowledgment

The authors would like to thank the management of Charutar Vidya Mandal, Vallabh Vidyanagar and management of A D Patel Institute of Technology for their continuous support and encouragement. The authors would also like to thank the Rsoft design group of OptSim5.2 simulation software and Fiber Optics Services, Bombay.

References

[1] E.Desurvire, *Erbium-Doped Fiber amplifiers – Principles and Application*, Wiley & Sons, 1st edition, 1994.

[2] P.C.Becker, N.A.Olsson, and J.R.Simpson, *Erbium-doped Fiber Amplifiers: fundamental and Technology*, San Diego, Academic Press, 1st edition, 1999.

[3] G.Keiser, *Optical Fiber communications*, McGraw-Hill International Editions, 3rd Edition, 2000.

[4] A.A.M.Saleh, R.M.Jopson, J.D.Evankow and J.Aspell, "Modeling of Gain in Erbium- Doped Fiber Amplifiers," *IEEE Photonics Technology Letters, vol. 2, pp. 714-717, 1990.*

[5] C. R. Giles and E. Desurvire, "Modeling Erbium-Doped Fiber Amplifiers," *Journal of Lightwave Technology, vol.9, pp. 271-283, 1991.*

[6] D.H.Richards, J.L.Jackel, M. A. Ali, "Multichannel EDFA Chain Control: A Comparison of Two All-Optical Approaches," *IEEE Photonics Technology Letters, vol. 10, pp. 156-158, 1998.*

[7] S.Milo, R.F.Souza, Marcia B.C.Silva, E.Conforti and A.C.Bordonalli, "An EDFA Theoretical Analysis Considering Different Configurations and Pumping Wavelengths," *IEEE IMOC, vol. 5, pp. 105-110, 2003*

[8] M. Achtenhagen, R.J.Besson, F. Pan, B. Nyman, A.Hardy, "Gain and Noise in Ytterbium-Sensitized Erbium-doped Fiber Amplifiers: Measurements and Simulations," *Journal of Lightwave Technology, vol.19, pp.1521-1526, 2001.*

[9] O. Lumholt, T.Rasmussen, A.Bjarlev, "Modelling of Extermely High Concentration Erbium Doped Silica Waveguides," *Electronics Letters, vol.29, pp. 495-496, 1993.*

[10] D.T.Nguyen, A.Chavez, S.Jiang and N.Peyghambarian, "A Novel Approach of modelling Cladding Pumped Highly Er-Yb co-doped Fiber Amplifiers," *IEEE Journal of Quantum Electronics, vol.43, pp. 1018-1027, 2007.*

Ronak R. Vashi received the BE degree from Birla Vishvakrarma Mahavidyalaya, Vallabh Vidyanagar, Gujarat, India. He is currently pursuing the ME in Communication Engineering from G.H.Patel College of Engineering and Technology, V.V. Nagar.
His research interests include Optical Fiber Communication, Optical Devices, Antenna and Microwave Systems.

Pravin R. Prajapati received the BE degree from Government Engineering College, Modasa, Gujarat, and the M Tech(Communication Systems) from Institute of Technology, Banaras Hindu University, Varanasi, India.
He is Associate Professor, Department of Electronics and Communication at A.D.Patel Institute of Technology, Karamsad, Gujarat, India.
His research interests include Optical Fiber Communication, Optical Devices, Power Electronics and Communication Systems.
Pravin Prajapati is life member of ISTE.

Effect of Solar flares on Background of X-ray Astronomical Objects

Nirav Y Pandya, Lecturer, *Electrical Engg. Department, Gandhinagar Institute of Technology*
nirav.pandya@git.org.in

ABSTRACT: - We know usually that various phenomenon takes place in Sun's atmosphere due to its activity. During the Solar flare: widely explosion in Sun's atmosphere; high speed electrons hurled out into planetary space where they affect satellites, earth's atmosphere and various astronomical objects. But the aim of my work is mainly to understand "How solar flares affect the background of X-ray astronomical objects."

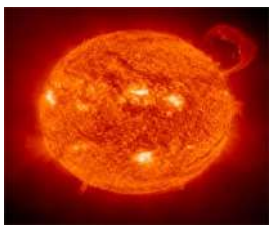
To study this effect, I have taken highest X-ray flux solar flare data from GOES-10 satellite and also taken the images of background of X-ray astronomical objects for the same time period. On the base of this data, I have drawn the graph of background count rate versus flux of selected X-ray flare which clearly indicates as X-ray flux increases the background count rate also increases. Thus X-ray flux solar flare increases the background of X-ray astronomical objects.

It is necessary to study this type of effects occur in the universe because if there exists very faint sources compared to the background surrounding X-ray sources, X-ray astronomical satellite cannot detect this type of faint sources and it can be go for false prediction about that sources.

INTRODUCTION

We know usually that various phenomena take place in the sun's atmosphere due to its activity. Solar flare is one such violent phenomenon. During the solar flare: widely explosion in the sun's atmosphere, high speed electrons hurled out into planetary space where they affect satellites, earth's atmosphere and various astronomical objects. The aim of my project is to understand how the solar flares affect the background during the X-ray astronomical observation.

SOLAR FLARE: BRIEF EXPLOSION IN THE SUN

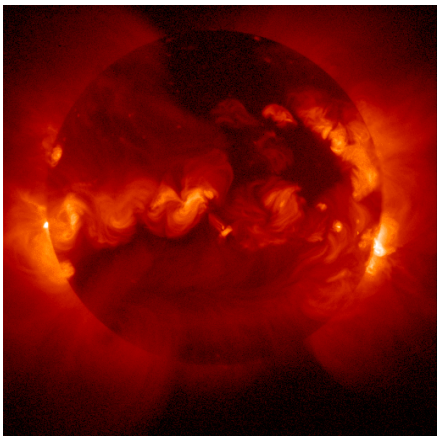


What is a solar flare?

A solar flare is defined as a sudden and brief explosion from the active region on the sun's surface with the energy equivalent to 2.5million terrestrial nuclear bombs, each with a destructive force of 100 megatons (10^{11} kg) of TNT (trinitrotoluene). A solar flare occurs when the magnetic energy that has built up in the solar atmosphere is suddenly released. Radiation is emitted across virtually the entire electromagnetic spectrum from radio waves at higher wavelength end to the x-ray and the gamma rays at the shorter wavelength.

X-ray solar flare

Since solar flares are very hot, they emit the bulk of their energy at x-ray wavelengths. For a short while large flare can outshine the entire sun in x-rays.



After a solar flare:

Shown here is an image of the Sun in soft x-rays. The white (brightest) region on the right hand side shows post-flare loops, hot loops that remain after a solar flare. Image from the Yohkoh Soft X-ray Telescope, from NASA's Observatories.

Solar flare can be characterized by their brightness in X-rays as observed by monitoring satellites near the earth. The biggest flares are X-class flare which has the flux of 10^-4 and above. M-class flares have one tenth the X-ray flux of an x-class one and the c-class has one hundredth of the x-class flux.

Effect of solar flare in the space

During the sudden and brief outbursts protons and electrons can be accelerated to nearly the speed of light. The high speed electrons and protons are also hurled out into interplanetary space where they can threaten astronauts and satellites. During release of flare energy there is also the coronal mass ejection take place which also affect the interplanetary space and earth's atmosphere.

The intense radiation from a solar flare travels to earth in eight minutes. As a result:

- The earth's upper atmosphere becomes more ionized and expands.
- Long distance radio signals can be disrupted by the resulting change in the ionosphere.
- a satellite orbit around the earth can be disturbed by the enhanced drag on the satellite from the expanded atmosphere.
- Satellite's electronic component can be damaged.

INTRODUCTION TO X-RAY ASTRONOMY

X-rays are a form of light, but much more energetic then light detected by our eyes. The energy of an x-ray photon is ~1000 times that of a photon of visible light. They are part of the electromagnetic spectrum which is shown below in the figure.

Where as astronomical x-ray is usually imaged in a different way from normal x-rays. In astronomy we image the source of x-ray itself. Sometimes however, there is an object which does get in the way and this then can appear as a shadow on the x-ray emission.

Generally the X-ray astronomical object has million degree K temperature and X-rays are emit from the hot object. So study of the astronomy at x-ray wavelength can give us more information about astronomical object. So X-ray astronomy was started.

Today the study of x-ray astronomy to be carried out using data from a host of satellites past and present, the HEAO series, EXOSAT, Ginga, CGRO,RXTE, ROSAT, ASCA. On 20th July 1999, the CHANDRA observatory launched by NASA and on December 10, 1999 the European x-ray satellite XMM (x-ray multi mirror mission) NEWTON was launched for x-ray observation.

A wide variety of x-ray sources have been known seen since identification of the first extra solar X-ray source in 1962. The Sun,Stars,Comets,X-ray Binaries, Supernova Remnates, Quasars and Active Galactic Nuclei are most interesting X-ray sources.

Roentgen used photographic plates to detect x-rays for his experiment. Now astronomers use more modern methods of detection. Today CCD (CHARGE COUPLED DEVICE) is used as x-ray detector.

A CCD is an image sensor, consisting of an integrated circuit containing an array of linked or coupled, light sensitive capacitors. The CCD was invented in 1969 by Willard Boyle and George E.Smith at AT&T labs. Astronomer for many years used CCD in ultraviolet bands, they are, however, new to x-ray astronomy with their first use on the Japanese-American satellite for ASCA (Advance satellite for cosmology and astrophysics) launched in 1992.

We know that X-rays are detected by earth's atmosphere and because of it the rocket flight which could left the payload in the earth's upper atmosphere were necessary. For this purpose the satellite was the good option and so X-ray astronomy satellite was developed for X-ray astronomy.

X-ray astronomy satellite: NASA's advanced x-ray astrophysics facility (AXAF), renamed the Chandra x-ray observatory in honor of Subramanian Chandrasekhar; XMM (X-ray Multi Mirror) Newton Observatory and also few new mission (ASTRO E2 "SUZAKU") were launched for the purpose of X-ray astronomy.

X-RAY SOLAR FLARES DATA FROM GOES

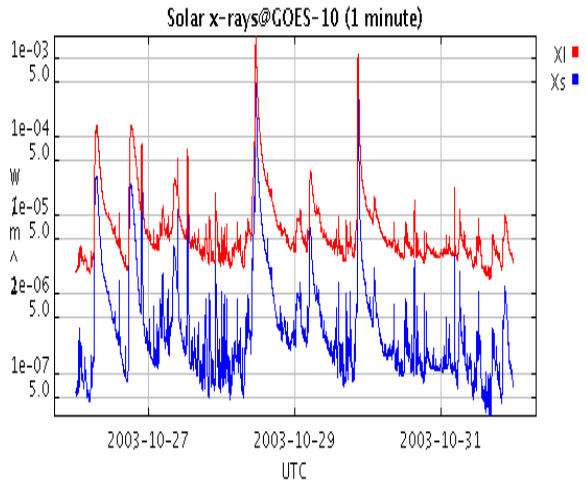
The Space Environmental Centre of the National oceanic and atmospheric administration, or NOAA provides the peak soft x-ray flux for solar flares seen from their Geostationary operational environmental satellites or GOES for short. The satellite hovers above the points in the earth's western hemisphere, orbiting at the same rate that that earth spins. I have taken the soft x-ray flux data for solar flare from GOES Space Environmental monitor.

There are various GOES satellite provides the data of the x-ray flux for the solar flare.

GOES-5 (Jan 1, 1986-Mar 31, 1987)
GOES -6(Jan 1, 1986-nov 30, 1984)
GOES-7 (Mar 1, 1995-June 30, 2003)
GOES-8 (Mar 1, 1995-June 30, 2003)
GOES-9 (Apr 1, 1996-Aug 31, 1998)
GOES-10(Jul 1, 1998-Apr30, 2007)
GOES-11 (Aug 1, 2000-Apr 30, 2007)
GOES-12 (Jan 1, 2003-Apr 30, 2007)

Here I have taken the data of the solar x-ray flux from 2000 to 2006 using the GOES-10 satellite. Here I want to find out the date and time on which the highest X-ray flux solar flare had occurred.

For it, I have taken the curve of the solar x-ray flux in w/m^2 Versus UT (universal time) for the duration of the every five days and note down the time and flux for the different x-ray flare. In this data I have found large number of M-class solar flare and few numbers of the highest x-ray solar flare i.e. X-class solar flare. Here I got approximately hundred X-class solar flare in which the highest solar flare occurred on 28 October, 2003 at 11:00 UT with the highest X-ray flux of 0.0025 w/ m^2. This solar flare had great effect on the solar system. The curve of X-ray flux versus UTC for this highest X-ray solar flare is shown below.



Here XL 1-8 Ang X-ray flux (watt/m^2)
XS 0.5-3 Ang X-ray flux (watt/ m^2)

In above fig. we can see that the first and third highest flux solar flare occur within a week. The first highest and third highest solar flare occurs on 28 October, 2003 a11:00 UT and 29 October, 2003 at 22:00 UT. Due to these two highest X-ray flux solar flare, there is tremendous effect can be seen on the earth's atmosphere, Which we will see later in result.

XMM-NEWTON DATA PREPARATION

The basic reduction of XMM-Newton data is provided by the SAS software package produced by the XMM- Newton Survey Science centre (SSC).

The SAS provides tasks to produced calibrated photon event files from the observation data format (ODF) files supplied to the user as well as some basic analysis tools. These include data filtering and extraction, production of rmfs and arfs for spectral analysis, and source detection routines.
The SAS was not designed for high level scientific analysis such as spectral fitting and temporal analysis. However, the SAS product files are designed to OGIP standards, so theoretically any high-energy is capable of processing XMM-Newton data. For example,

o HEASoft from the High Energy Astrophysics Science Archive Research Centre (HEASARC)
o CIAO from the Chandra X-ray observatory centre (CXC) can both be used for XMM-Newton data analysis.

The high energy astrophysics science archive research centre (HEASARC) is the primary archive for high-energy astronomy missions, in the extreme ultraviolet, X-ray and gamma-ray wavelengths. The HEASARC provides archival data, multi mission software and analysis tools, and information about current and past observatory missions.

We can download XMM-Newton data on a CD-ROM or from the public archive at VILSPA or also from HEASARC. HEASARC provides the data for all the observatories, data analysis tools etc.

I have download XMM-Newton data for randomly selection of 15 highest X-class solar flare. For ex: Highest solar flare occurs on 28 October, 2003 so I have download XMM-Newton data for the duration of one day before to six days after means from 27 october,2003 to 03 November,2003. There are various type of data available in HEASARC browse archive for the XMM-Newton data. Among them I have taken it from XMMmaster (XMM master log and public archive).

Once if we got our XMM observation data, we can choose it between the observation data file (ODF) and the already processed product files created by an automatic pipeline. The advantage of the pipeline products is that you do not have to create the event files etc. yourself on your machine, which sometimes takes several hours and you can start immediately working with the scientifically data. The ODF data contains all of the observation specific data necessary for reprocessing the observation. The pipeline data contain among other things calibrated photon event files, source lists and images. Here I have used only pipeline product data.

XMM –NEWTON DATA ANALYSIS

EXAMINE AND FILTER AN EPIC DATA - PIPELINE PRODUCTS

The EPIC event lists in the EEVLIS group of the Pipeline Processing will have names of the form:

- PiiiiiijjkkaaSlllcIEVLI0000.FTZ, where

  iiiiiijjkk - observation number
  aa - detector (M1 - MOS1, M2 - MOS2, PN - PN)
  lll - exposure number within the observation
  c - detector (M - MOS1 or MOS2, P - PN, T - Timing Mode)

The following sections describe the use of SAS tasks using the both the command-line and GUI interfaces, except in cases where one of the methods is particularly easy. The SAS GUI provides a very simple method for producing and displaying images, spectra, and light curves, and is the recommended method for extracting data unless large numbers of sources are being analysis.

Initialize SAS and Prepare the Data

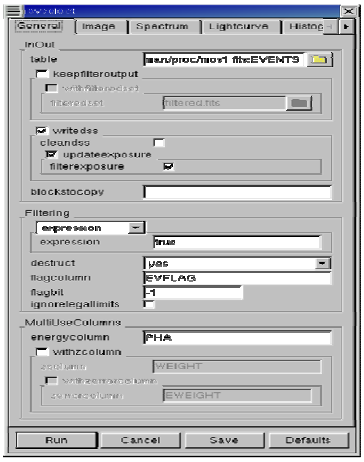
(1) Gunzip the PP event list to be examined (not really necessary), and for practical purposes shorten the file name as well, e.g.:
mv P0123700101M1S001MIEVLI0000.FTZ mos1.fits.gz
gunzip mos1.fits.gz
(2) For initializing SAS GUI
•sasinit
(3) Invoke the sas GUI (in below fig.1)

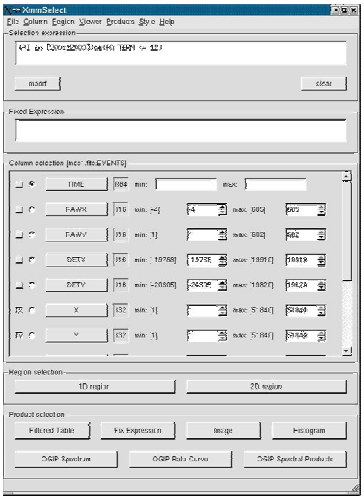
(4) Invoke the xmmselect GUI from the SAS GUI; double click on the task name.

When here xmmselect is invoked a dialog box will first appear requesting a file name. After double click on the desired event files in the right hand column, click on file in the “EVENT” extension in the right hand column and then click on “OK.” The directory GUI will then disappear and then click “RUN” on the selection GUI.

When the filename has been submitted the xmmselect GUI (fig 2) GUI will appear along with a dialog box offering to display the selection expression will include the filtering done to this point on the event file.

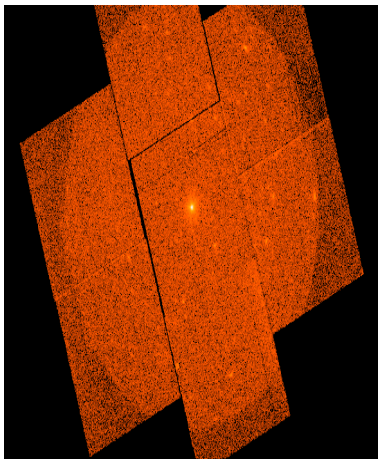
In xmmselect GUI we can choose the two dimensional data selecting the square boxes on the left hand side. In this case X,Y sky coordinates have been selected while one dimensional data can be selected using the round boxes.

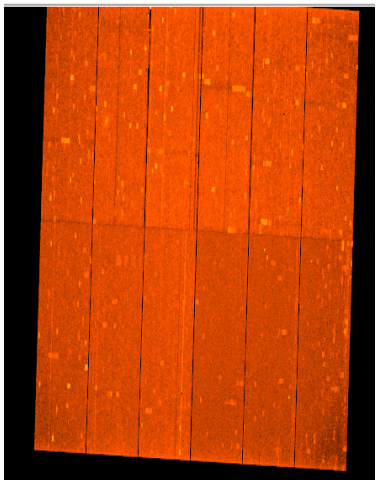




Create and Display an Image

- To create an image of the data in sky coordinates check the square boxes to the left of the ``X" and ``Y" entries.
- Click on the ``Image" button near the bottom of the page. This brings up the evselect GUI.
- The default settings are reasonable for a basic image so click on the ``Run" button at the lower left corner of the evselect GUI. Different binnings and other selections can be invoked by accessing the ``Image" tab at the top of the GUI.
- The resultant image is written to the file `image.ds`, and the image is automatically displayed using ds9.

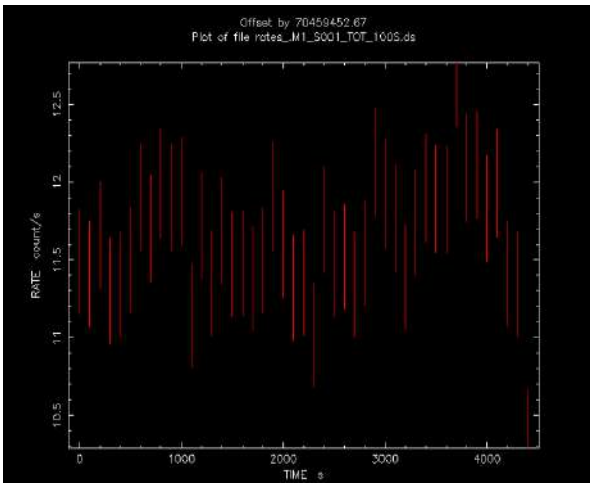




Two different images for EPIC MOS1 and in EPIC PN are shown in above fig. the left side fig. is for EPIC MOS with seven CCD and right side is for EPIC PN with 12 CCD.

Create and Display a Light Curve

- To create a light curve check the round box to the left of the ``Time" entry.
- Click on the ``OGIP Rate Curve" button near the bottom of the page. This brings up the evselect GUI.
- The default setting is for a one-second bin which is a bit fine, so access the ``Light curve" tab and change the ``timebinsize" to, e.g., 100 (100 s). Click on the ``Run" button at the lower left corner of the evselect GUI.
- The resultant light curve is written to the file rates.ds, and is displayed automatically using Grace.

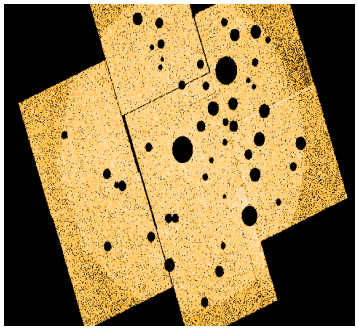


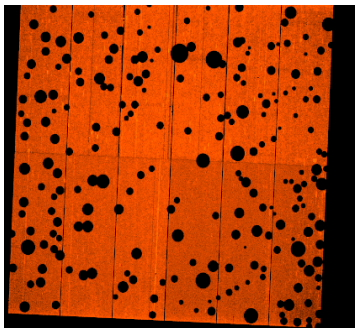
Above fig. shows the total light curve with the X-ray source for MOS1 which indicates that count rate for peak flare is 12.8 c/s on time 3700s for 100s background light curve.

Filter the Data and Create a New Event File

Filter the data using the xmmselect GUI: Here purpose of filtering is to get the background image of the MOS1 data and light curve for this background images. For it, in the image source and we want to extract the sources and we want to keep only the background area of the image.

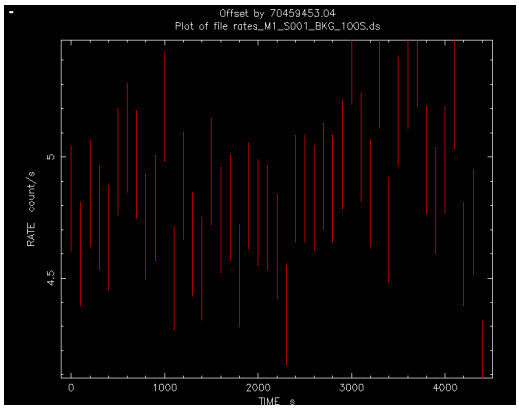
For it, we select the each region belong to the xmmselect and then click 2D region. Doing in this way we can get the selection expression for the selected source region and then to extract it, we put "!" sign before it. So we can get only background region in the images. Then click on IMAGE and giving particular name for the background image. Thus we obtain background image. Two different background images for EPIC MOS and EPIC PN are shown as below.





Create and display a Background light curve

Here we got the selection expression for the background region of the images. Now in xmmselect click on OGIP rate curve and after giving particular name to it and taking particular time (here I took 100s,10s and 1s) for each light curve and then click on enter we can get the background light curve for the particular time. The background light curve for this type of selection expression in which the x-ray source region are extracted is shown in below fig.



Average background count rate for 100s light curve

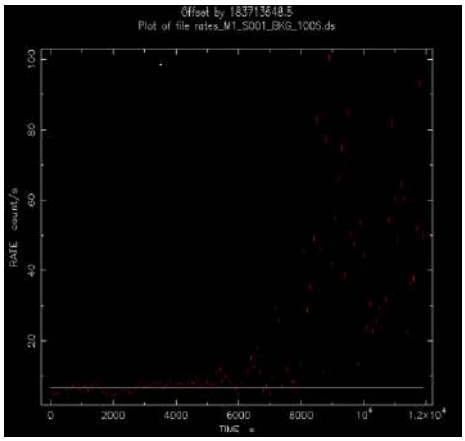
As shown further we got background light curve for particular time and instrument containing certain observation I.D. I have used the command fplot, mo. (model), co.(constant) and then fit for find out the avg. background count rate for base level for each 100s light curve. Then I have plot the graph of avg. background count rate versus the range of the X-ray flux for

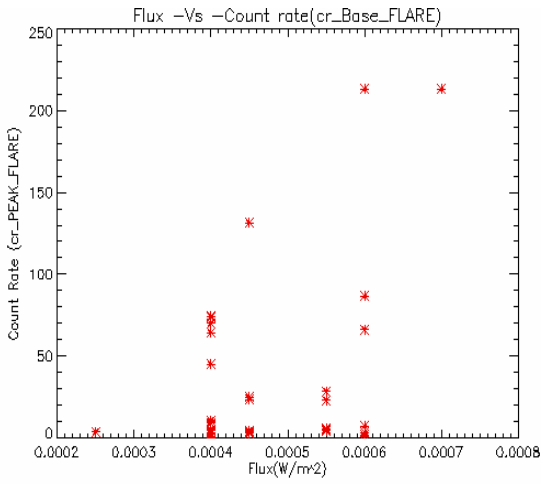
selected flare. From it I have got important result which is shown below.

RESULTS

● There are two important plots, which are shown below, describe the important result of my work. The second plot is main result of my project which is for the combined observation of both MOS1 and MOS2. It is the graph for background count rate versus flux of the selected X-ray flare. Here we can clearly see that as the flux increases the background also increases. It is not very clear result but we can see that background increases with the flux of X-ray solar flare.

● Now in the second fig., the background light curve for the duration of the highest X-ray solar flare which clearly indicates that the background clearly increases after the highest X-ray flux solar flare. Thus we can see that the emission of the outbursts during the solar flare, increase the background of the X-ray astronomical object.





● Here during my project work I have observed the interesting effect of the solar flare on the XMM-Newton satellite. Before the highest X-ray flux solar flare occurs on 28, October 2003 the XMM-Newton was taking the observations but after the highest solar flare, I couldn't take my XMM-Newton data because it stops to take the observation and again it starts to take the observation on 9 November, 2003. Thus I observed the tremendous effect of it in the satellite.

CONCLUSIONS

From my all work on this project the "effect of solar activity mainly solar flare on the background of X-ray astronomical object.", I conclude that due to the outbursts going outside the sun, the background of the astronomical object increases.

It is necessary to study the this type of effect occurs in the universe because if there exists very faint sources compared to the background surrounding them and due to this, the X-ray astronomy satellite can't observe this type of sources and it can be go for false prediction about that sources. So it is very necessary to study this type of effect which occurs in the universe.

FUTURE ASPECTS

- Here I have taken the XMM-Newton data for sparsely selected highest X-class solar flare. If we take XMM-Newton data for all the X-class Solar flares, we can study the effect of solar flare on the background of the X-ray astronomical object more completely.
- I have studied the background light curve only for the 100s duration in this project. If we study the background light curve for 1s duration, then we can get understand effect of the solar flare on the temporal studies Of the X-ray astronomical objects.
- Similarly if we include the spectral analysis method, we can get more important information about the effect of solar flare on the X-ray Spectral studies of the X-ray astronomical objects.

REFERENCES

- The restless universe: understanding X-ray astronomy in the age of Chandra and Newton by Eric M. Schlegel, Oxford University.
- The Cambridge encyclopedia of the sun: Kenneth R. Lang
- GOES data archive: http://goes.ngdc.noaa.gov
- HEASARC browse archive: http://heasarc.nasa.gov/W3Browse
- XMM-Newton ABC Guide: http://heasarc.gsfc.nasa.gov/docs/xmm/abc/abc.html
- XMM –Newton data cookbook: http://wave.xray.mpe.mpg.de/xmm/cookbook

Analysis of Self-describing Gridded Geo-science Data Using High Performance Cluster Computing

Ms. Kruti Lavingia
Information Technology Dept.,
Sabar Institute of Tech. for Girls, Tajpur
E-mail: krutilavingia@gmail.com

Mr. Nilay Shah
Information Technology Dept.,
E-mail: nilay@nilayshah.in

***Abstract*— Unidata's netCDF (network Common Data Form) data model, data access libraries, and machine independent format are widely used in the creation, access, and sharing of geoscience data. NetCDF is widely used in earth, ocean, and atmospheric sciences because of its simple data model, ease of use, portability, and strong user support infrastructure.**

The netCDF format provides a platform-independent binary representation for self-describing data in a form that permits efficient access to a small subset of a large data set, without first reading through all the preceding data. The format also allows appending data along one dimension without copying the data set or redefining its structure.

But a geoscience researcher cannot deal with programming environment, so to develop such an interface which can make reading and writing netCDF data possible for them is the scope of the project.

The interface needs to be implemented on high performance computing platforms to deal with large data files which results into long delays or out of memory problem while serial execution. Java is to be used as High Performance Computation Language and JOMP for providing parallel programming interface.

***Index terms*—netCDF, JOMP, Cluster**

I. INTRODUCTION

1.1 General

NetCDF (network Common Data Form) is a set of software libraries and machine independent data formats that support the creation, access, and sharing of array-oriented scientific data.NetCDF (network Common Data Form) is a set of interfaces for array-oriented data access and a freely-distributed collection of data access libraries for C, Fortran, C++, Java, and other languages. The netCDF libraries support a machine-independent format for representing scientific data. Together, the interfaces, libraries, and format support the creation, access, and sharing of scientific data. NetCDF Data is: Self-Describing. A netCDF file includes information about the data it contains. Portable. A netCDF file can be accessed by computers with different ways of storing integers, characters, and floating-point numbers. Direct-access : A small subset of a large dataset may be accessed efficiently, without first reading through all the preceding data. Appendable : Data may be appended to a properly structured netCDF file without copying the dataset or redefining its structure. Sharable: One writer and multiple readers may simultaneously access the same netCDF file. Archievable: Access to all earlier forms of netCDF data will be supported by current and future versions of the software.

1.1.1 Why NetCDF:

Unidata's netCDF (network Common Data Form) is a data model for array-oriented scientific data access, a package of freely available software that implements the data model, and a machine-independent data format. NetCDF supports the creation, manipulation, and sharing of scientific data sets that are self-describing, portable, directly accessible, and appendable. A data model specifies data components, relationships, and operations, independent of any particular programming language. The components of a netCDF data set are its variables, dimensions, and attributes. Each variable has a name, a shape determined by its dimensions, a type, some attributes, and values. Variable attributes represent ancillary

information, such as units and special values used for missing data. Operations on netCDF components include creation, renaming, inquiring, writing, and reading. The netCDF software includes interfaces for C, Fortran, C++, perl, and Java. Utilities are available for displaying the structure and contents of a netCDF data set, as well as for generating a netCDF data set from a simple text representation. The netCDF format provides a platform-independent binary representation for self-describing data in a form that permits efficient access to a small subset of a large data set, without first reading through all the preceding data. The format also allows appending data along one dimension without copying the data set or redefining its structure. Since Unidata developed netCDF, other groups and projects in the geosciences have adopted the netCDF interfaces and format, and its use has also spread to other disciplines.

1.1.2 NetCDF Features:

- Multiple unlimited dimensions
- Portable structured types
- String type
- Additional numeric types
- Variable-length types for ragged arrays
- Unicode names
- Eficient dynamic schema changes
- Multidimensional tiling (chunking)
- Per variable compression
- Parallel I/O
- Nested scopes using Groups

1.2 Motivation

Geoscience Data Files are available to geoscience researchers from various sources, but reading, visualization and writing a file for them is not easier. Although various software tools are available for viewing netCDF file data, but then also reading particular variable, then modification in original data and then writing netCDF file using high performance cluster computing is an important task for any research scientist which is not available in any tool. This project will develop an interface for geoscience researchers, by which they can do all the geoscience data analysis tasks.

1.3 Scope Of Work

The scope of work starts with Study of NetCDF file data model and all the task related to accessing the file via programming interface. Choosing a particular programming language as Java and then, developing an interface to read, write, modify netCDF files. The interface should be able to read and edit any variables, attributes and data of any netCDF file. Finding the best possible way to implement parallel programming interface using language Java. And implementation of the interface on High Performance Cluster Computing using JOMP for providing parallel programming interface.

II. NetCDF

2.1 NetCDF 4 Architecture:

NetCDF-4 uses HDF5 for storage, high performance.

- Parallel I/O.
- Chunking for efficient access in different orders.
- Conversion using "reader makes right approach".

Provides simple netCDF interface to subset of HDF5. Also supports netCDF classic and 64-bit formats. netCDF-4 Architecture is shown in figure 2.1.

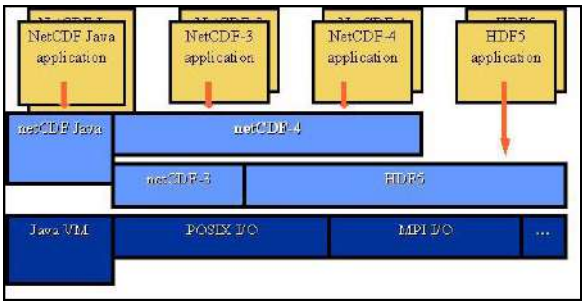


Figure 2.1: NetCDF-4 Architecture

2.2 NetCDF Usage

Since netCDF was made available in 1989, the popularity of the interface and format has continued to grow. Now widely used in the atmospheric sciences, it is one of only a handful of data-access interfaces and formats that are used across diverse scientific disciplines. For example, as part of the Distributed Ocean Data System (DODS), developers have implemented a client-server-based distributed system for access to oceanographic data over the Internet that supports use of the netCDF interface for clients. Descriptions of some of other projects and groups that are now using netCDF are available from Unidata. NetCDF data may now be accessed from over 20 packages of freely available software, including DDI, DODS, EPIC,

FAN, FERRET, GMT, GrADS, HDF interface, LinkWinds, SciAn, and Zebra. Access to netCDF data is also available from commercial or licensed packages for data analysis and visualization, including IBM Data Explorer, IDL, GEMPAK, MATLAB, PPLUS, PV-Wave, PolyPaint+, and NCAR Graphics. The unexpectedly widespread use of netCDF means that any future changes to the data model, interfaces, or format must be planned and implemented with great care. Backward compatibility with existing software and data archives is very important to netCDF users and must be part of future development plans.

2.3 Benefits of netCDF

Benefits of using netCDF or other similar higher-level data-access interfaces for portable and self-describing data include :

- Sharing common data files among different applications, written in different languages, running on different computer architectures;
- Reduction of programming effort spent interpreting application- or machine-specific formats;
- Incorporation of metadata with the data, reducing possibilities for misinterpreting the data;
- Accessing small subsets of data efficiently from large data sets;
- Making programs immune to changes caused by the addition of new variables or other additions to the data schema; and
- Raising the level of data issues to structure and content rather than format.

2.4 The Binary Formats

By "binary formats" we mean the layout of bytes on the disk. NetCDF-4.0 supports three binary data formats:

- classic - the original netCDF binary data format
- 64-bit offset - the variant format which allows for much larger data files
- netCDF-4 - the HDF5-based format, with netCDF-specific constraints and conventions.

Additionally there is one "virtual" format: netCDF-4 classic model. This format
is obtained by passing the classic model ag when creating the netCDF-4 data file. Such a file will use the netCDF-4 format restricted to the classic netCDF data model. Such files can be accessed by existing programs that are linked to the netCDF-4 library.

The Programming APIs and Libraries

- By "programming APIs and Libraries" we mean the software that makes netCDF available in various computer programming languages.
- The language APIs are implemented in two distinct core libraries: the original C library and the independent Java library. The Fortran and C++ APIs call the C library functions. All other APIs not in a Java environment are based on the C library.
- NetCDF-4 has been fully implemented in the C library; implementation in the Java library is underway.

2.5 The Data Model

By "data model" we mean the way scientific data is conceptually modeled with a
set of objects, operations, and rules that determine how the data is represented and accessed. The classic model, as shown in figure 2.2, of netCDF represents data as a set of multi-dimensional arrays, with sharable dimensions, and additional metadata attached to individual arrays or the entire file. In netCDF terminology, the data arrays are variables, which may share dimensions, and may have attached attributes. Attributes may also be attached to the file as a whole. One dimension may be of unlimited length, so data may be efficiently appended to variables along that dimension. Variables and attributes have one of six primitive data types: char, byte, short, int, oat, or double.

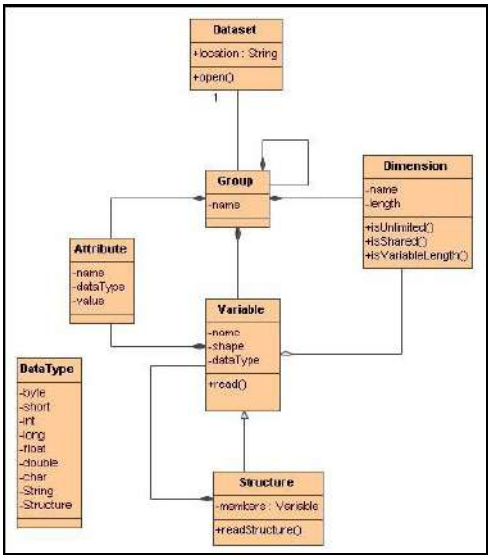


Figure 2.2: NetCDF-4 Data Model

NetCDF-4 expands this model to include elements from the HDF5 data model, including

hierarchical grouping, additional primitive data types, and user defined data types. The new data model is a superset of the existing data model. With the addition of a nameless "root group" in every netCDF file, the classic model fits within the netCDF-4 model.

2.6 Using NetCDF4 with the Classic Data Model

NetCDF-4 brings many new features to users within the classic netCDF model. By confining themselves to the classic model, data producers ensure that their data files can be read by any existing netCDF software which has been re linked with the netCDF-4 library. For example, the use of a compound type in a file requires the netCDF-4 data model, but reading compressed data does not. One advantage of only using features that conform to the classic data model is that existing code that reads, analyzes, or visualizes the data will continue to work. No code changes are needed for such programs, and they can transparently use netCDF-4 features such as large file and object sizes, compression, control of endianness, reading chunked data, and parallel I/O, without modification of existing code. For example, data producers can use zlib compression when writing out data files. Since this is transparent to the reader, the programs that read the data do not need to be modified to expand the data. That happens without any help from the reader. In many cases, users may wish to use netCDF-4 data files without adding any of the model-expanding features. As a convenience netCDF-4 includes the CLASSIC-MODEL flag. When a file is created with this flag, the rules of the classic netCDF model are strictly enforced in that file. This remains a property of the file, and the file may never contain user-defined types, groups, or any other objects that are not part of the classic netCDF data model.

Large File and Object Size: NetCDF-4 files may contain larger objects than classic netCDF or even 64-bit offset netCDF files. For example, variables that do not use the unlimited dimension cannot be larger than about 4 Giga Bytes in 64-bit offset netCDF files, but there is no such limit with netCDF-4 files on 64- bit platforms**.**

Compression and Shuffle Filters: NetCDF-4 uses the zlib library to allow data to be compressed and uncompressed as it is written and read. The data writer must set the appropriate ags, and the data will be compressed as it is written. Data readers do not have to be aware that the data are compressed, because the expansion of the data as it read is completely transparent. The shuffle filter does not compress the data, but may assist with the compression of integer data. The shuffle algorithm changes the byte order in the data stream; when used with integers that are all close together, this results in a better compression ratio. There is no benefit from using the shuffle filter without also using compression. Data compression and shuffling may be set on a per variable basis. That is, the zlib compression eg (from 0,no compression, to 9, maximum compression) can be set independently for each variable. In our tests we notice that setting the deate higher than one takes more time, but has little benefit**.**

Control of Endianness: In netCDF classic format files (and 64-bit offset format files), numeric data are stored in big-endian format. On little-endian platforms, netCDF is converted to big-endian when the data are written, and converted back to little endian when read from the file. In netCDF-4 files, the user has direct control over the endianness of the each data variable. The default is to write the data in the native endianness of the machine. This is useful in cases where the data are to be read on the same machine, or machines of similar architecture. However, in some cases the data may be produced on a machine of one native endianness, and read on a machine of the other endianness. In these cases, the data writer may wish to optimize for the reader by explicitly setting the endianness of the variable. In our tests, the endianness of the data only affected read rates significantly when disk caches were in full use, and the data were read from the disk cache. In this case, data with a native endianness were read noticeably faster. However, when disks caches were cleared, the endianness of the data does not affect the read rate much. Apparently the disk speed is slow enough without caching that the CPU has plenty of time to swap the bytes of the data while waiting for the disk. When the data are available in cache, the I/O rate is much faster, and then the cost of the byte swapping becomes noticeable. For high-performance applications in which netCDF file reading is a bottleneck and access patterns allow disk caching to be used effectively, users should consider writing variables in the file with the endianness of the target platform. Higher-performance disk systems may also serve the data fast enough for its endianness to matter.

Chunking: NetCDF-4 files may be written as chunked data, each chunk representing a multidimensional tile of the same size. That is, the data are written as chunks of a given size, specified by the user when the variable is created and before any data is written. Compressed variables must be chunked, and each chunk is compressed or uncompressed independently. Chunking has important performance ramifications. Both file size and I/O rates are affected by chunk sizes, and choosing very small chunk sizes can be disastrous for performance. The following graph shows the file sizes of the radar 2D sample data for a variety of chunk sizes. Chunk sizes should be chosen to yield an amount of data that can be comfortably handled by disk buffers. Chunk sizes that are too small or too large result in poor performance or overly large data files. Since compression and expansion work on individual chunks, specifying too large a chunk size may cause a large portion of a file to be uncompressed when reading only a small subset of the data. One heuristic for data providers to use is square chunks about one megabyte in size. Chunk sizes should also be chosen so that a whole number multiple of the chunk completely fills the dimension. Users will also experience better performance by using contiguous storage for variables of fixed size, if data are accessed sequentially.

Parallel I/O: NetCDF-4 supports parallel I/O on platforms that support MPI (the Message Passing library). Parallel I/O in netCDF-4 only works on netCDF-4 data files.NetCDF-4 users may use special functions to open or create files, to which they can write data in parallel, and from which they can read data in parallel. Parallel data reads can result in significant performance improvements in some high-performance computing applications. Equivalent wrapper functions for the Fortran APIs are provided in the netCDF distribution. Recent testing on TeraGrid machines showed clear performance gains with parallel I/O, on parallel file systems with low processor counts.

III. A COMMON DATA ACCESS MODEL FOR GEOSCIENCE DATA

An effort to provide useful mappings among NetCDF, HDF, and OpeNDAP data abstractions as shown in figure 2.3 Intended to enhance interoperability. Let scientists do science instead of data management.

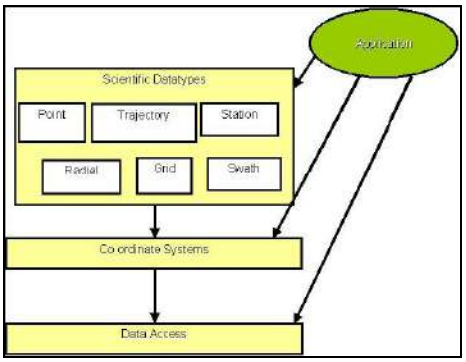


Figure 2.3: A Common Data Access Model for Geoscience Data

Lets data providers and application developers work more independently. Raises level of discourse about data objects, conventions, coordinate systems, and data management. Demonstrated in NetCDF-Java 2.2, which can access netCDF, HDF5, OpeNDAP, GRIB1, GRIB2, NEXRAD, NIDS, DORADE, DMSP, GINI, ... data through a single interface! NetCDF-4.0 C interface implements data access layer.

IV. USING JAVA AS HIGH PERFORMANCE COMPUTING LANGUAGE

Java offers a number of benefits as a language for High Performance Computing (HPC). For example, Java offers a high level of platform independence not observed with traditional HPC languages. This is an advantage in an area where the lifetime of application codes exceeds that of most machines. In addition, the object-oriented nature of Java facilitates code re-use and reduces development time. However, there are a number of outstanding issues surrounding the use of Java for HPC, principally: performance, numerical concerns and lack of standardized parallel programming models. There is a wide variety of interfaces and language extensions for parallel and distributed programming in Java. Both [9, Java threads] and Remote Method Invocation (RMI) are part of the Java specification. [9, Java threads], although principally designed for concurrent, rather than parallel, programming can successfully be used on shared memory multiprocessors. RMI is not well suited to parallel programming, both due to its programming paradigm and its high overheads. The two interfaces which we have used are MPJ and JOMP. These are prototype specifications of

Java counterparts to MPI and [10, OpenMP] respectively. We have chosen these interfaces due to the familiarity and widespread use of their Fortran and C predecessors, and the fact that neither requires extension to the core Java language. It should be noted that neither is yet standardized, and so may be subject to change in the future. [2, MPJ], as shown in figure 4.1, consists of a class library providing an interface for message passing, similar to the MPI interface for C and Fortran. Most of the functionality found in MPI is supported, and messages may consist of arrays of either basic types or of objects. Existing implementations such as [8, mpiJava] use the Java Native Interface (JNI) mechanism to call existing MPI libraries written in C. However, research efforts are underway to provide pure Java implementations using sockets or VIA. [3, JOMP], as shown in figure 4.2, is a specification of directives (embedded in standard Java as comments), runtime properties and a class library similar to the OpenMP interface for C and Fortran. The existing implementation uses a source-to-source translator (itself written in Java) to convert the directives to calls to a runtime library, which in turns uses the standard [9, Java threads] interface. The system is pure Java, and therefore transparently portable. Other approaches to providing parallel programming interfaces for Java include JavaParty, [1, HPJava] , Titanium and SPAR Java. These are also in the research phase and, in addition, require genuine language extensions.

V. CONCLUSION

The application interface for accessing geoscientific file and analyzing its data has been developed which can help geoscience researchers to easily interact with it. For development of the application using high performance computing, Java is used as high performance computation language and JOMP is used for providing Parallel Programming Interface. I have used an OpenMP-like interface for Java which enables a high level approach to shared memory parallel programming. A prototype compiler and runtime library which implement most of the interface have been described, showing that the approach is feasible. Only minor changes from the OpenMP C/C++ specification are required, and the implementation of both the runtime library and the compiler are shown to be relatively straightforward. Initial analysis shows that the resulting code scales well, with little overhead compared to a hand-coded Java threads version. Low-level synchronization overheads have been measured and are for the most part, tolerable. For data intensive cluster, execution time for parallel execution will be less than serial execution. Loading variable data and modification of data become faster than before.

REFERENCES

[1] G. Fox X. Li B. Carpenter, G. Zhang and Y. Wen. Hpjava: Data parallel
extensions to java. In Concurrency: Practice and Experience, 1998.

[2] M.A. Baker and D.B. Carpenter. Mpj: A proposed java message-passing api and environment for high performance computing. Second Java Workshop at IPDPS, Cancun, Mexico, LNCS, Springer Verlag, Heidelberg, Germany, pages 552,559, 2000.

[3] J. M. Bull and M. E. Kambites. Jomp an openmp-like interface for java. In The ACM 2000 Java Grande Conference, pages 44,53, June 2000.

[4] Edward Hartnett and R. K. Rew. Experience with an enhanced netcdf data model and interface for scienti_c data access. UCAR, Boulder, CO.

[5] M. D. Westhead D. S. Henty J. M. Bull, L. A. Smith and R. A. Davey. A methodology for benchmarking java grande applications. In ACM 1999 Java Grande Conference, pages 81,88, 1999.

[6] Mark Kambites. Java openmp.

[7] W. Liao A. Choudhary R. Ross R. Thakur W. Gropp R. Latham A. Siegel B. Gallagher M. Zingale Li, J. Parallel netcdf: A high-performance scienti_c i/o interface. ACM, 2003.

[8] G. Fox S.-H. Ko M. Baker, B. Carpenter and S Lim. mpijava: An object- oriented java interface to mpi. In International Workshop on Java for Parallel and Distributed Computing, IPPS/SPDP, April 1999.

[9] S. Oaks and H. Wong. Java Threads. O'Reilly.

[10] Openmp architecture review board, October 1998.

Microelectromechanical Systems (MEMS) devices, Sensing , Actuation and Applications

Neelu Kanabar
Department of Instrumentation & Control
Sabar Institute of Technology for Girls
At & Po. Tajpur,
India
Neelam_kanabar@yahoo.com

Binal Gandhi
Department of Instrumentation & Control
Sabar Institute of Technology for Girls
At & Po. Tajpur,
India
Sagi.bini@gmail.com

***Abstract*— Microelectromechanical systems (MEMS) integrate mechanical and electrical components and have feature sizes ranging from micrometers to millimeters. They may be fabricated using methods similar to those used to construct integrated circuits and they have the potential of providing significant cost advantages when batch fabricated. Their size also makes it possible to integrate them into a wide range of systems. Feature sizes may be made with size of the order of the wavelength of light, thus making them attractive for many optical applications. Micro sensors (e.g., accelerometers for automobile crash detection and pressure sensors for biomedical applications) and micro actuators (e.g., for moving arrays of micro mirrors in projection systems) are examples of commercial applications of MEMS**

Keywords- MEMS, fabrication, micrometer, smart vehicle

I. INTRODUCTION

A micro electro mechanical system (MEMS) is the integration of mechanical elements, sensors, actuators, and electronics on a common silicon substrate through micro fabrication technology MEMS are micrometer-scale devices that integrate electrical and mechanical elements. They have been used in diverse applications, from display technologies to sensor systems to optical networks. MEMS are attractive for many applications because of their small size and weight, which allow systems to be miniaturized.

II. MEMS TECHNOLOGY DRIVERS (ADVANTAGES) AND CHALLENGES:

Key technology drivers:

Cost: The primary driver for the initial adoption of MEMS technology was cost. MEMS technology has a huge potential for bringing the costs down, as it can be easily batch fabricated. For example, Analog Devices Inc.'s fabrication method has standardized so as to ship out millions of accelerometers every year. The process is streamlined and requires less labor compared to conventional manufacturing of sensors. Furthermore, many MEMS fabrication technologies allow parallel fabrication of thousands of systems by leveraging the parallel fabrication techniques of the integrated circuit industry. This may lead to a reduction in the manufacturing cost and improvement in reliability.

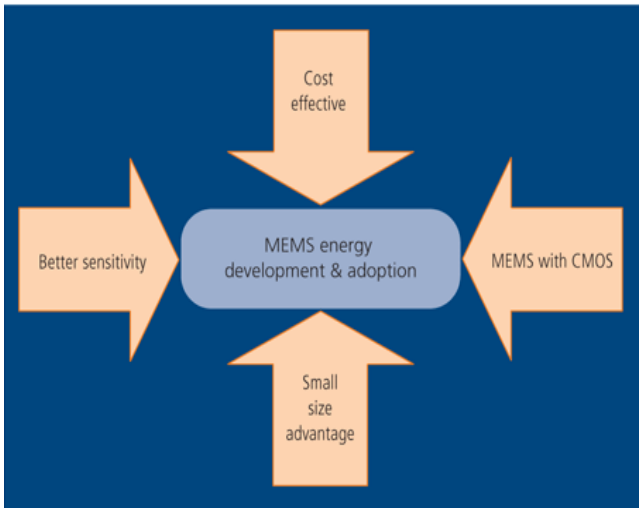


Figure 1: Key technology drivers for MEMS

The small size of MEMS is attractive for many applications because feature sizes are typically as small as 1 micrometer or less. Hence, for optical applications, features may be made with size on the order of the wavelength of light. Their small size also allows applications to be developed which would otherwise be impossible. For example, micromechanical switches fabricated as part of a communications circuit allow phase shifting and signal switching at speeds that would be impossible to achieve using macro-scale switches. To illustrate the scale of a typical micro system, Figure 2 below shows a micro machined mirror assembly next to a spider mite. The mirror assembly is about 100 micrometer wide, and it is dwarfed by the spider mite. The mechanical devices surrounding the mirror allow it to be positioned accurately as part of an optical network.

Increased sensitivity: MEMS sensors can be more sensitive compared to conventional sensors. This is a major reason

MEMS sensors are preferred over conventional sensors, especially in applications that can benefit from high accuracy and precision.

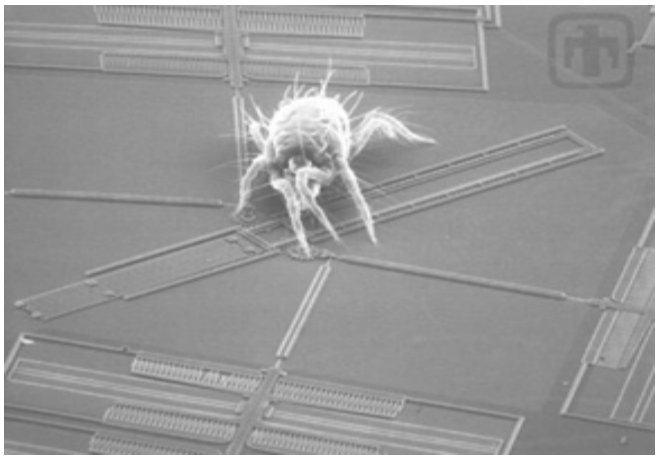


Figure 2: A spider mite dwarfs a micro machined mirror assembly. Courtesy Sandia National Laboratories, SUMMiT Technologies, www.mems.sandia.gov.

Other advantages include the on-chip integration of electromechanical systems and the circuitry used to control them, allowing further miniaturization

Challenges:

Like any technology, micro systems present some challenges.

1. Because micro mechanisms operate at a size scale far below that of typical mechanical devices, surface forces such as adhesion and friction may dominate over other forces in the system, leading to failure of the device. Careful device design, fabrication, and testing are required to reduce these effects.

2. Their small size also makes it difficult to interact mechanically with MEMS components. In many micro electromechanical systems, electrical or optical signals are used to interface with, provide power to, and control the device, rather than manual, hydraulic, or pneumatic control typically seen in macro-scale mechanical devices.

3. In addition, packaging of MEMS components has often presented a challenge because each device must be packaged in a way that keeps the components clean and free from contamination, while also allowing mechanical motion and, in many cases, interaction with the environment. For example, a MEMS pressure sensor requires a package that exposes the sensor to the ambient pressure while protecting the electronic circuitry from dust or other particles.

4. Finally, while parallel fabrication techniques can reduce the manufacturing cost of many units, MEMS development may be more costly because few units are produced at a time using complex and expensive fabrication equipment.

5. MEMS technology is also application specific, and hence a very small number of standardized MEMS processes exist. For instance, pressure sensors, accelerometers, ink-jet printers, displays, and all other varieties of MEMS employ different process modules. Hence, there is a lack of synergy or cooperation in design, packaging, testing, and tool development.

III. FABRICATION TECHNOLOGIES

A large variety of fabrication methods have been employed for MEMS. However, many of these methods may be broadly described under three headings: surface micromachining, bulk micromachining and processes with high aspect ratios (molding) like DRIE and LIGA

A. Bulk Micromachining

Bulk micromachining, the oldest of the micromachining technologies, is accomplished by removing material from a substrate to create holes, cavities, channels, or other desired shapes.

B. Surface Micromachining

Surface micromachining is one of the most common technologies used to manufacture MEMS devices. In surface micromachining, films are deposited on a substrate and patterned, using photolithography, to create micromechanical devices. The films normally alternate between structural and sacrificial layers, with the MEMS parts being made from the structural layers. The sacrificial layers serve to support the structural components during fabrication. After the structural layers are patterned, the sacrificial material is removed, often using wet chemical etching. The result is freestanding MEMS parts that can move relative to the fixed substrate.

C. Comparison of Micro fabrication Processes

Each of the processes discussed above has advantages and disadvantages, depending on the application. Several are summarized in the table below. Table 1 describes the relative complexity of parts which are normally created using each process, the typical relative size of the MEMS resulting from each process, the complexity of the fabrication steps required, and the typical aspect ratio of the parts, meaning the ratio of their height to their width. The table is very general, but provides a rule-of-thumb comparison of the different processes.

	Part Complexity	Typical Size	Fabrication Complexity	Aspect Ratio
Surface Micromachining	High	Small	Low-High	Low
Bulk Micromachining	Low	Large	Low	High
Molding(LIGA)	Medium-High	Small-Large	Medium	High

TABLE 1: COMPARISON OF MICRO FABRICATION PROCESSES

IV. MEMS ACTUATION AND SENSING METHODS:

A. Actuation

On-chip actuation of micro systems has been a particularly challenging aspect of MEMS development. Common macro-level actuation approaches, such as hydraulics, pneumatics, electric motors, internal combustion engines and turbines, are either too difficult to fabricate at the micro level or do not work well at that scale. Electrostatic attraction is one approach that has been widely used for actuation of Microsystems. While electrostatic actuation is suitable for many applications, some systems require either lower voltages or higher output forces.

1) Electrostatic Actuation

According to Coulomb's law, the electrostatic force acting between two charges is inversely proportional to the distance between the charges. For macro-scale objects, this force is normally negligible. However, micro-scale devices may have very small gaps, making electrostatic attraction an important source of mechanical motion fig 3. This actuation technology is especially attractive because it uses very little power. On the other hand, large voltages (typically tens to hundreds of volts) are required.

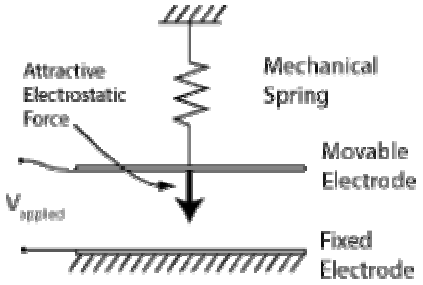


Figure3: A parallel-plate actuator consists of two parallel electrodes. Attractive electrostatic force pulls the movable electrode toward the fixed electrode**.**

B. Sensors

A sensor is a device that responds to a physical input (such as motion, radiation, heat, pressure, magnetic field), and transmits a resulting signal that is usually used for detection, measurement, or control. A transducer (often used as a synonym for sensor) is a device that is actuated by power from one system and converts it to a different form to another system. Advantages of MEMS sensors are their size and their ability to be more closely integrated with their associated electronics.

Piezoresistive and capacitive sensing methods are among the most commonly employed sensing methods in MEMS. Piezoresistance is the change in resistivity caused by mechanical stresses applied to a material. Materials with high piezoresistivity (such as some semiconductors which have more than an order of magnitude higher piezoresistivity than metals) are useful for transducing mechanical deformation to electrical signals. This is particularly useful in applications such as pressure sensors and accelerometers.

Capacitive sensors rely on the physical input being sensed to cause a change in capacitance. This capacitance change can be caused by changing the distance between the capacitor plates (e.g. pressure pushing two plates closer together) or by changing the dielectric (such as relative humidity sensor using a dielecric with a permittivity that changes with moisture content). The resulting change in capacitance can be very small and specialized electronics are required to detect the changes and convert them into a usable output signal. Mechanical transducers

V. MEMS APPLICATION EXAMPLES

1. Mechanical transducers:
   Strain gages, Accelerometers, Gyroscopes, Pressure sensors, Microphones, Electrostatic motor, Shape Memory Alloy, Piezoelectric, Mechanical resonators, Mechanical Relays and RF switches
2. Optical transducers:
   Photo conductive sensors, Junction-based photo detectors, Capacitive photo sensors, Thermal Optical detector, Light Emitters, Reflective and Transmissive Micromechanical light Modulators, Fiber optic couplers, Reflective components (mirror), Transmissive components (Wave guide, lenses), Filters, Integrated Optical systems
3. Magnetic transducers:
   Hall effect sensor, Flux gate, Tunneling, Motors, Eddy current detector, Magnetic read/write head, Micro-Inductors
4. Thermal transducers:
   Thermal Resistive, Thermal Couple, Junction based Thermal sensors, Thermal Gas pressure sensor, Flow sensor, Humidity sensor and Peltier effect heat pumps
5. Chemical and Biological transducers:
   Passive Chemical Sensors, Electrochemical Transducers, Biosensors, Biological Chemical Sensors (taste Odor), Thin Film Batteries, Penetrating Neural Probes
6. Fluid devices:
   Flow Channel, Mixer, Pumps, Valve, Separator, Droplet generators, Filters, Interconnects, Flow sensors, Viscosity/Density Sensors, Ink-Jet printer heads

VI. APPLICATION OF MICROSYSTEMS IN AUTOMOTIVE INDUSTRY

The automotive industry has been the major user of MEMS technology in the last two decades because of the size of its market The primary motivation for adopting MEMS and Microsystems in automobiles is to make automobiles safer and more comfortable for riders and to meet the high fuel efficiencies and low emissions standards required by governments The term *smart cars* was first introduced in the cover story of a special issue of a trade magazine (Smart Cars, 1988). Many of the seemingly fictitious predictions of the intelligent functions of a smart car are in place in today's vehicles.

Smart vehicles are built on the extensive use of sensors and actuators. Various kinds of sensors are used to detect the environment or road conditions, and the actuators are used to execute whatever actions are required to deal with these conditions. Microsensors and actuators allow automobile makers to use smaller devices, and thus more of them, to cope with the situation in much more effective ways. .Figure4 illustrates the application of pressure sensors in an automobile.

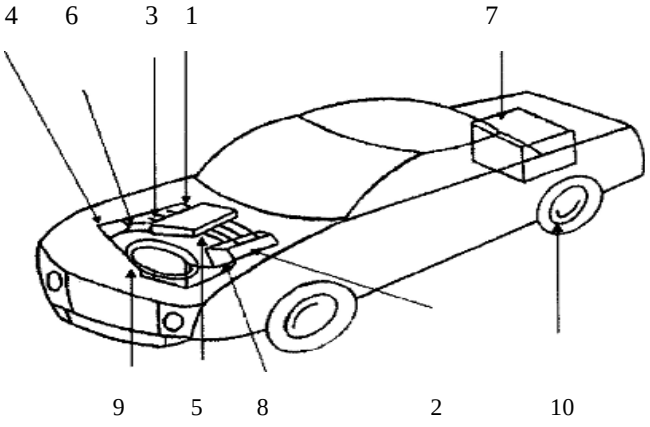


Figure 4. Pressure sensors in automotive applications.

(1) Manifold or Temperature manifold
Absolute pressure sensor
(2) Exhaust gas differential pressure sensor
(3) Fuel rail pressure sensor
(4) Barometric absolute pressure sensor
(5) Combustion sensor
(7) Fuel tank evaporative fuel pressure sensor
(6) Gasoline direct injection pressure sensor
(8) Engine oil sensor
(9) Transmission sensor
(10) Tire pressure sensor

Applications of Microsystems in automobiles can be categorized into the following four major areas:
(1) Safety,
(2) Engine and power train,
(3) Comfort and convenience,
(4) Vehicle diagnostics and health monitoring.

VII. CONCLUSION

The concept of integrating more than one transistor on a single chip has had a profound and lasting impact on our society. From the first use in the Minuteman missile to the proliferation of consumer products today, microelectronic circuits have dramatically improved performance, functionality, and reliability, while reducing cost and decreasing volume.

Today researchers pursue advanced concepts that integrate not only electronics at the micro scale, but embody sensors, photonics and MEMS components. MEMS offer opportunities to miniaturize devices, integrate them with electronics, and realize cost savings through batch fabrication.

MEMS technology has enhanced many important applications not only in industrial automation but also in medical and more recently in consumer electronics and it holds great promise for continued contributions in the future. The real explosion for MEMS could be just over the horizon

REFERENCES

[1] An Introduction to Microelectromechanical Systems Engineering – N. Maluf
[2] MEMS mechanical sensors – Stephen Beeby
[3] MEMS – Nitaigour Premchand Mahalik (SPIE Press ,second reprint 2008)
[4] Sensors and Microsystems – P. Siciliano, S. Capone, L. Di Natale
[5] MEMS AND MOEMS tech and applications –P. Rai Choudhary
[6] Dsp design line (tech online community)
[7] www.isa.org/link/200401 (mems magic)
[8] www.csa.com (oct 2001 Salvator Vittorio)
[9] www.memx.com
[10] www.mems.sandia.gov

Performance Analysis of Antenna Diversity Techniques: A Way to Improve the Bit Error Rate of WiMAX System

Nirali A. Kotak [1], Bhavin S. Sedani [2]

[1] Department of Electronics & Communication Engineering,
Gandhinagar Institute of Technology, Ghandhinagar, Gujarat, India
E-mail: nirali.kotak@git.org.in Ph: +91 9925200505

[2] Department of Electronics & Communication Engineering,
VVP Engineering College of Technology – Vajdi Virda, Rajkot - Gujarat, India
E-mail: bhavin_s_sedani@yahoo.com Ph: +91 9925041418

Abstract

Today life does not seem possible without wireless in some form or the other. The demands on bandwidth and spectral availability are endless. gConsequently, wireless designers face an uphill task of limited availability of radio frequency spectrum and complex time varying problems in the wireless channel, such as fading and multipath, as well as meeting the demand for high data rates.

Wireless communication is one of the most vibrant areas in the communication field today. In the last few years, the telecommunication industries development has focused on an intensive use of broadband systems, with high quality features. For this issue, new technologies with high transmission abilities have been designed. There is one fundamental aspects of wireless communication that make the problem challenging and interesting, that is the phenomenon of fading.

Traditionally the design of wireless systems has been focused on increasing the reliability of the air interface where fading and interference are viewed as nuisances. Recent focus is now on spectral efficiency; associated with this shift is a new point of view that fading can be viewed as an opportunity to be exploited. In the initial stages, the single antenna system was used at transmitter and receiver in both the sides (Single Input Single Output – SISO system) anticipating lower spectral efficiency and lesser capacity. It can be seen that root cause of the poor performance of these techniques is that reliable communication depends on the strength of the single signal path only which may be in a deep fade under certain circumstances. A natural solution to improve the performance is to ensure that the information symbols pass through multiple signal paths, each of which fades independently, making sure that reliable communication is possible as long as one of the paths is strong. This technique is called diversity, and it can dramatically improve the performance over fading channels.

The main objective of this paper is to achieve the greater system capacity with reduction in bit error rate for wireless system by implementing various antenna diversity principles (Singe Input Multiple Output – SIMO, Multiple Input Single Output - MISO & Multiple Input Multiple Output - MIMO) along with space time coding and multiple antennas either at the transmitter or receiver or at both the sides and to analyze the same features for WiMAX system.

1. Overview of WiMAX:

The experienced growth in the use of digital networks has led to the need for the design of new communication networks with higher capacity. The increased reliance on computer networking and the Internet has resulted in a wider demand for connectivity to be provided "any where, any time", leading to a rise in **the requirements for higher capacity** and high reliability broadband wireless telecommunication systems. **WiMAX may be seen as the fourth generation (4G) of mobile systems** as the convergence of cellular telephony, computing, Internet access, and potentially many multimedia applications become a real fact.

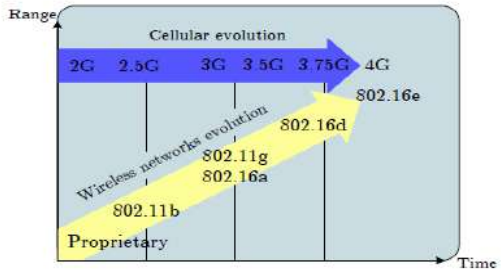


Figure-1 Cellular evolution and Wireless network evolution

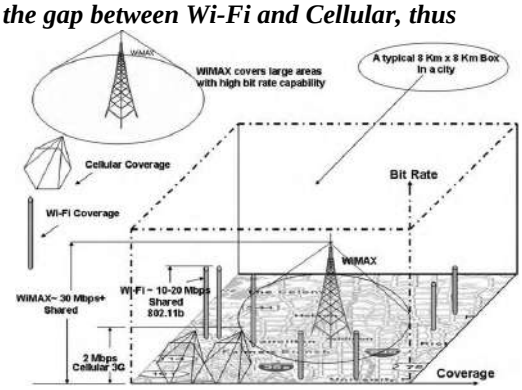
WiMAX (Worldwide Interoperability for Microwave Access) is considered today the most interesting opportunity, able to provide radio coverage distances of almost 50 kilometers and data troughput up to 70 Mbps, and to complete wired network architectures, ensuring a flexible and cheap solution for the last-mile**. The interoperability** is a very strategic issue, on which equipment cost and volume of sales will be based. Operators will not be bound to a unique equipment supplier, as the radio base stations will be able to interact with terminals produced by different suppliers. Operators can benefit of suppliers' competition in terms of costs and innovation.

2. Relation ship with other wireless technology:

Wireless access to data networks is expected to be an area of rapid growth for mobile communication systems. The huge uptake rate of mobile phone technologies, WLANs and the exponential growth that is experiencing the use of the Internet have resulted in an increased demand for new methods to obtain high capacity wireless networks.
WiMAX is expected to have an explosive growth, as well as the WiFi, but compared with the Wi-Fi WiMAX provides broadband connections in greater areas, measured in square kilometers, even with links not in line of sight. For these reasons WiMAX is a MAN, highlighting that "metropolitan" is referred to the extension of the areas and not to the density of population. But WiFi and WiMAX are not competing technologies. While WiMAX can provide high capacity internet access to residences and business seats, WiFi allows the extension of such connections inside the corporate sites buildings.

In any case, both WLAN and cellular mobile applications are being widely expanded to offer the demanded wireless access. However, they experience several difficulties for reaching a complete mobile broadband access, bounded by factors such as bandwidth, coverage area, and infrastructure costs.
As shown in above figure, **Wi-Fi provides a high data rate, but only on a short range of distances and with a slow movement of the user**. On the other hand, **Cellular offers larger ranges and vehicular mobility, but instead, it provides lower data rates, and requires high investments**
for its deployment.
WiMAX tries to balance this situation. ***WiMAX fills the gap between Wi-Fi and Cellular, thus***

Figure-2



Comparative analysis

providing vehicular mobility, and high service areas and data rates.

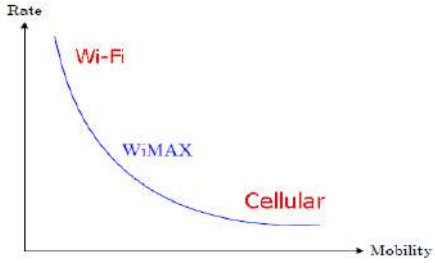


Figure-3 Relationship with other wireless Technology.

3. WiMAX -OFDM

The WiMAX network technology is an evolutionary one as it uses orthogonal frequency-division multiplexing (OFDM), which makes the transmissions resistant to fades and multipath effects.
In real-life transmission environments, multipath propagation and echoes from objects lead to the received signals arriving at the destination in a time-delayed fashion. These signals suffer frequency selective fading as a result of the multipath propagation effects. When single carrier is used to carry high data rates (i.e., a short symbol time), the received signals have enough delay spread to fall into the slots of other symbols thereby causing intersymbol interference.
In the case of a single carrier modulation, this type of propagation limits the data rates that can be used in non-line-of-sight (NLOS) environments. **The technology of OFDM is based on the use of a large number of carriers spread in the allocated bandwidth, with each carrier being modulated by a proportionately lower data rate than would be the case in a single-carrier scenario**.

OFDM can be viewed as a form of frequency division multiplexing (FDM) with the special property that each tone is orthogonal with every other tone, but it is different from FDM in several ways. On one hand, **FDM requires, typically, the existence of frequency guard bands between the frequencies so that they do not interfere with each other.** On the other hand, **OFDM allows the spectrum of each tone to overlap, and because they are orthogonal, they do not interfere with each other.** Furthermore, the overall amount of required spectrum is reduced due to the overlapping of the tones.

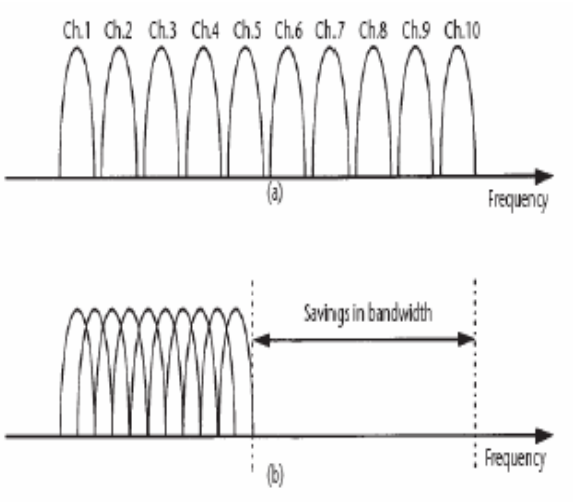


Figure-4 orthogonal frequency-division multiplexing

In OFDM systems, each carrier is orthogonal with respect to the other carriers, so that the sidebands of the carrier cancel out rather than interfering with the next carrier. The principle of using orthogonal carriers has given the name OFDM to the technology of using multiple carriers. OFDM is a common technology used in achieving high

bit rates in all wireless and wireline systems, which may be subject to frequency selective fading or intersymbol interference from reflected signals. OFDM systems provide a very robust transmission technique for NLOS environments. OFDM is also used in wire line systems such as DSL, and will be used in the next generation of the long-term evolution of cellular technologies.

4. Antenna systems and diversity:

WiMAX defines a number of optional advanced features for improving the performance. Significant gains in overall system capacity and spectral efficiency can be achieved by deploying the optional advanced antenna systems (AAS) defined in WiMAX.
Among the more important of these advanced features are support for multiple-antenna techniques, hybrid-ARQ, and enhanced frequency reuse.

4.1 Single Input Single Output (SISO)

The main fundamental behind advance antenna system implementation is the diversity. In the initially stages, the various modulation schemes like coherent BPSK, coherent QPSK, coherent 4-PAM, coherent 16-QAM were in which error probability decay very slowly proportional to 1/SNR.

Basically the above mention modulation techniques did not use diversity principle that is the single antenna system were used at transmitter and receiver in both the side anticipating lower poor spectral efficiency and lesser capacity. It can be seen that root cause of the poor performance of these techniques is that reliable communication depends on the strength of the single signal path only. There is a significance probability that these path will be in a deep fade under certain circumstances.

When the path is in a deep fade, any communication scheme will likely suffer from errors. A natural solution to improve the performance is to ensure that the information symbols pass through multiple signal paths, each of which fades independently, making sure that reliable communication is possible as long as one of the paths is strong. This technique is called Diversity, and it can dramatically improve the performance over fading channels.

4.2 Types of Diversity scheme

There are three ways to obtain diversity.

1) Time Diversity
2) Frequency Diversity
3) Space Diversity

Diversity over time can be obtained via coding and interleaving: information is coded and the coded symbols are dispersed over time in different coherence periods so that different parts of the code words experience independent fades. Analogously, one can also exploit diversity over frequency if the channel is frequency-selective. In a channel with multiple transmit or receive antennas spaced sufficiently far enough, diversity can be obtained over **space** as well.

Three kinds of space diversity can be obtained to improve the capacity of the system.

1) Tx-Diversity in which multiple transmit antennae are used for the signal transmission which in term results in Multiple Input Single Output diversity (n x 1 system).

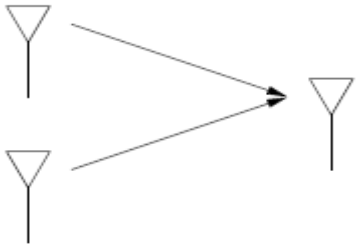


Figure-5 MISO system

2) Rx-Diversity in which multiple receive antennae are used for the signal reception which in term results in Single Input Multiple Output (1 x n system).

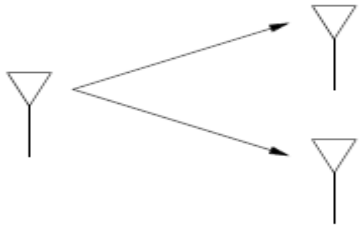


Figure-6 SIMO system

3) Tx-Rx Diversity in which Channels with Multiple Transmit and Multiple Receive antennas (So called Multi Input Multi Output or **MIMO** channels) provide even more potentional. In addition to provide diversity, MIMO channels also provide additional degree of freedom for communication.

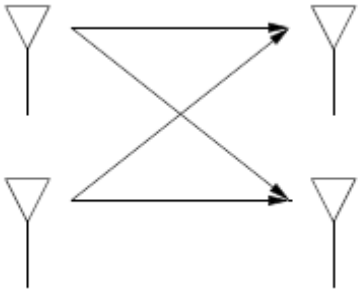


Figure-7 MIMO system

The main advantages of **MIMO** channels over **SISO** channels are

- The array gain,
- The diversity gain, and
- The multiplexing gain.

Array gain and diversity gain are not exclusive of MIMO channels and also exist in SIMO and MISO channels. Multiplexing gain, however, is a unique characteristic of MIMO channels. Array gain is the improvement in SINR obtained by coherently combining the signals on multiple transmit or multiple receive dimensions and is easily characterized as a shift of the BER curve due to the gain in SINR. Diversity gain is the improvement in link reliability obtained by receiving replicas of the information signal through independently fading links,

branches, or dimensions. It is characterized by a steepen slope of the BER curve in the low BER region.

5. Space-Time Coding (Alamouti Coding):

Transmit diversity makes use of space-time block coding schemes to provide transmit diversity in the downlink. This essentially requires at least two transmit antennas and at least one receive antenna. One of the choices of codes used is the Alamouti codes. The purpose of this scheme is not to increase the system capacity but to improve the error rate performance of the system by transmitting coded information.

The 802.16 standard i.e WiMAX specifies Alamouti scheme as a compliant STC scheme. It is a Space-Time code in that it sends information on two transmit antennas (space) over two consecutive transmissions in time. Therefore it is said to transmit information in space and time.

As shown in the following figure, the data stream entering the Modulator is modulated into Symbols S1 and S2. These symbols are then processed by the Space-Time Encoder which then sends S1 followed by –S2* to Antenna 1 and S2 followed by S1* to Antenna 2. Here (*) denotes a complex conjugate of the symbol. Note that the two antennas at the BS will transmit 2 symbols in two time periods. In other words this is a rate 1 code.

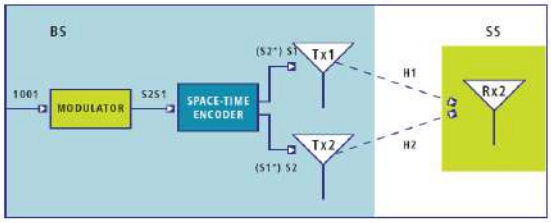


Figure 8 – STC Scheme

6. Results:

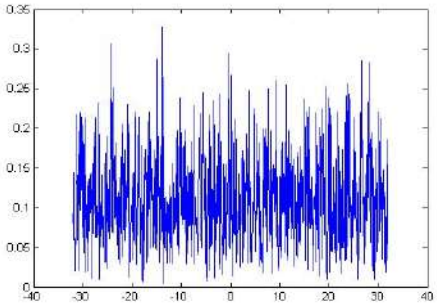


Figure-9 OFDM waves in time domain

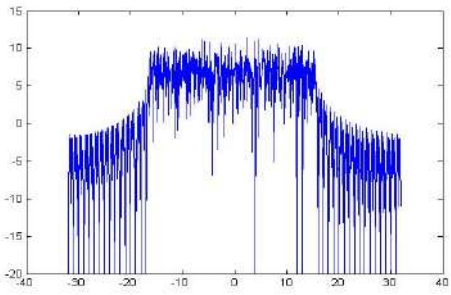


Figure-10 OFDM waves in frequency domain

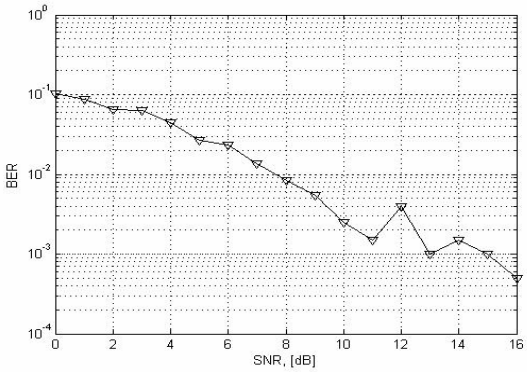


Figure-11 BER performance with 1 Transmit and 1 Receive Antenna (SISO)

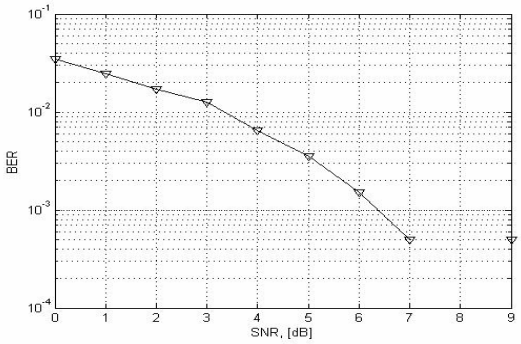


Figure-12 BER performance of Transmit Diversity with 1Transmit and 2 Receive Antennas

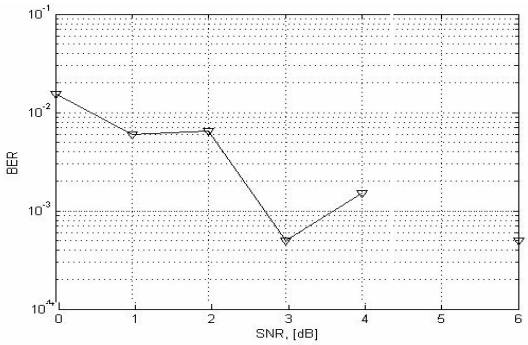


Figure-13 BER performance of Tx-Rx Diversity with 3Transmit and 3 Receive Antennas (MIMO)

6. Comments:

Simulations were performed to generate OFDM signals using Matlab. The plot of OFDM signal as well as the spectrum of OFDM is shown in figure by taking the below parameters.

Total Data symbol	512
Number of symbol per frame	32
Length of symbol for IFFT	64
Guard Interval length	8
Symbol Rate	1Hz
4 types of symbols	0,1,2,3
Symbol with GI Insertion	144

Table 1 Parameters to generate an OFDM signal

Also as shown in figure-8(c) and (d), the various curves of the BER v/s SNR under different diversity conditions shows that the performance of the system drastically improves with a diversity system.

As can be observed from the curve, the system performance in terms of bit error rate can be improved by adopting the fundamental of antenna diversity (SIMO and MIMO) as compared to 1 transmit 1 receive antenna system (no diversity).
Also the implementation of MIMO along with the Alamouti coding, the higher degree of diversity can be achieved which in turn increases the system throughput with low bit error rate.

7. Conclusion:

The key problem with the wireless channels is impairment of the channel by fading and interference. The major goal and the purpose of this paper is to analyze the new techniques which fulfills the increasing requirement of data rate and quality of service which increase spectrum efficiency and improve link quality. OFDM has proved to be very effective in mitigating adverse multipath effects of a broadband wireless channel. Multiple Input Multiple Output (MIMO) technique has proved its potential by increasing the link capacity significantly via spatial multiplexing and improving the link capacity via space time coding. Further it can be concluded that implementation of diversity technique is the way to improve the system performance in terms of bit error rate in WiMAX system.

8. References:

[1] David Tse, University of California, Berkeley Pramod Viswanath, Fundamental of Wireless Communication, published by Cambridge University Press, 2004

[2] H. Farhat, G. Grunfelder, A. Carcelen and G. El Zein," MIMO Channel Sounder at 3.5 GHz: Application to WiMAX System",JOURNAL COMMUNICATIONS, VOL. 3, NO. 5, OCTOBER 2008

[3] Jelena Mi¡si and Vojislav B. Misic, Wireless Personal Area Networks Performance Interconnections and Security with IEEE 802.15., John Wiley & Sons, Ltd

[4] Amitabh Kumar, "Mobile Broadcasting with WiMAX: Principles, Technology, and Applications", Focal Press Media Technology Professional Elsevier-2008

[5]Theodore S. Rappaport, "Wireless Communication Principle Practice", Second Edition, Printice Hall of India Private Limit 2006

[6] Hui Liu Guoqing Li "OFDM-Based Broadband Wireless Networks Design and Optimization", A JOHN WILEY & SONS, MC., PUBLICATION, 2005

[7] Allen W. Scott, Rex Frobenius "RF Measurements for Cellular Phones and Wireless Data Systems" Wiley-IEEE Press, 2008-07-15

[8] H. Farhat, G. El Zein, "Antenna arrays design and calibration for high resolution MIMO channel sounding at 3.5 GHz," In Proc. IEEE VTC Fall 2007, pp. 936-940,Baltimore, USA, Sept. 2007

[9] V. Kolmonen, J. Kivinen, L. Vuokko, P.Vainikainen, "5.3-GHz MIMO Radio Channel Sounder", IEEE Trans. On Instrum. Measur., Vol. 55, No. 4, pp. 1263-1269, Aug 2006

[10] http://www.wimaxforum.org/home/

[11] Onsy Abdel Alim1, Nemat Elboghdadly, Mahmoud M. Ashour, Azza M. Elaskary, "System-Level Model for OFDM WiMAX Transceiver in Radiation Environment" The Second All African IRPA Regional Radiation Protection Congress 22-26 April 2007 Ismailia Egypt.

[12] William C. Y. Lee, "Wireless & Cellular Telecommunications", McGraw-Hill Date : 2005

Author Biographies:

1. **Nirali A. Kotak** graduated in Electronics and Communication engineering from Atmiya Institute of Technology and Science, Rajkot in 2004. She has completed her post graduation in 2010 in the field of Electronics and Communication Engineering from CCET, Wadhwan. She is working as an Assistant Professor in Gandhinagar Institute of Technology, Gandhinagar, Gujarat in Electronics and Communication department. She has published several papers in international and national journals and conferences based on WiMAX, Antenna diversity, OFDM and Wireless channel modeling.

2. Bhavin S. Sedani graduated in Electronics and Communication engineering from C.U.Shah College of Engg. & Technology, Wadhwan in 2002. He has completed his post graduation in 2008 in the field of Electronics and Communication Engineering from CIT, Changa. He is pursuing Ph.D. from J.J.T. University, India. He is working as an Assistant Professor in V.V.P. Engg. College, Rajkot, Gujarat in Electronics and Communication department. He has published several papers in international and national journals and conferences based on WiMAX, Antenna diversity, OFDM and Wireless channel modeling.

Seismic Assessment of RC Frame Building with Masonry Infill Wall

PATEL J. J.[1], RAIJIWALA D. B.[2], AND DESAI M. K.[3]

[1] Civil Engineering Department, GIT, Moti Bhoyan, Gandhinagar, Gujarat.
[2&3] Applied Mechanics Department, SVNIT, Surat, Gujarat.

***Abstract*—Masonry walls in frame buildings are generally considered as nonstructural elements and their contribution in resisting lateral loads has therefore been neglected in the design of frame type structures. The presence of masonry infill walls are remarkable in increasing initial stiffness and strength of reinforced concrete (RC) frame buildings, and being the stiffer component, attract most of lateral seismic forces acting on buildings, thereby, reducing demand on RC frame members. However, buildings are often constructed with open first-storey for parking or other purposes. This introduces severe mass, stiffness, and strength irregularities in the open first-storey of these buildings. . In the following study selected building models are analyzed by non-linear equivalent static method in SAP2000 software. The infill represented by multiple diagonal struts,**

***Index Terms*— Multiple strut, Non linear static analysis, Opening, Soft storey, Stiffness irregularity**

I. INTRODUCTION

In multi-storey buildings the skeleton frames are filled in their plane by brick or concrete block as a wall to meet architectural and functional requirements, like partitioning, enclosing space etc. such walls which are not considered to carry any loads and which serve only non-structural purpose are called infilled walls, filler walls, partitions or filler panels. In such situations the combination of frames and filler panels form a composite structure called infilled frame, the combination leads to increasing the strength and lateral stiffness of the structure.

In normal practice, design engineers tend to treat the infill wall as non-structural element and design the frame as bare frames. However, it remains fact that, the frame and infill act together and generally offer combined resistances to applied loads. The advantageous behaviour of an infilled frame is derived mainly from the interaction that develops between the panels and the frame. In that;

- The in-bracing action of panels greatly increases the stiffness and strength of the infilled frames, as compared with the bare frame.
- The progressive yielding of the interface between the frame and the panels retain ductility of the structure.
- The progressive cracking of the panels dissipates the energy when excessive vibration occurs.

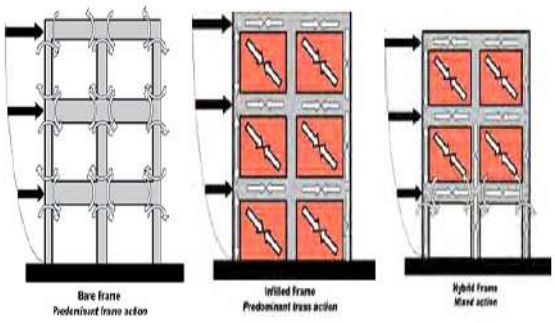


Fig. 1 Infill wall influence the behaviour of RC frame

If infill walls are not symmetrically placed in RC frames in plan, torsional effects are induced in the structure because of significant eccentricity introduced between centre of mass and centre of rigidity. Torsion introduce additional shear forces and bending moments in RC frame members for which these members are generally not designed for, leading to possible failure of some of these members. In several buildings, masonry infill walls are deliberately constructed only up to a particular height to serve various purposes, like saving the material, providing ventilation, architectural requirements, etc. Such construction practices create short columns in the storeys with partial height infill walls. Short columns tend to attract large shear forces than otherwise designed for and eventually such columns may fail in shear.

The interaction between the frame and the infill was found to be critical in the performance of the frame joints. Out-of-plane failures of the walls too were observed on many occasions. Opening in infill walls reduce their effectiveness in carrying lateral loads as the diagonal strut are not formed between the frame corners and thereby the overall truss action is not achieved in the building structures. Buildings with infill having large openings, also suffered significant local damage to the frame and out-of-plane collapses of those infills. The damage to Maruhong building in 1978 Miyagi-Ken Oki (Japan) earthquake was one of these kinds.

In 2001 Bhuj-earthquake, many modern reinforced concrete multi-storey buildings in Bhuj and Ahmedabad had collapsed. Amongst the multi-storey buildings that collapsed, most had

the ground storey left open for parking convenience with few or no filler walls between the columns. This created a top stiff inverted pendulum structure with insufficient strength and stiffness in the open ground storey, thereby rendering the same vulnerable.

The results in terms of natural period, storey displacement, base shear, storey stiffness, bending moment and axial force are compared for all models. The main objective of the present study is to understand the seismic behaviour of multi-storey RC frame building considering the effect of masonry infill wall with and without opening and soft storey.

II. MODELING OF MASONRY INFILL WALL

Experimental work indicated that when an infilled frame is subjected to lateral loads, transfer of load takes place through a truss action in the infill and this led to the development of equivalent diagonal strut models in which infills are replaced by single or multiple compression diagonal struts along the loaded diagonals.

A. Equivalent Single Strut:

The simplest equivalent strut model included a single pin-jointed strut with its width taken as one-third the diagonal length of infill, and connected between loaded diagonal of the frame [Holmes 1961].

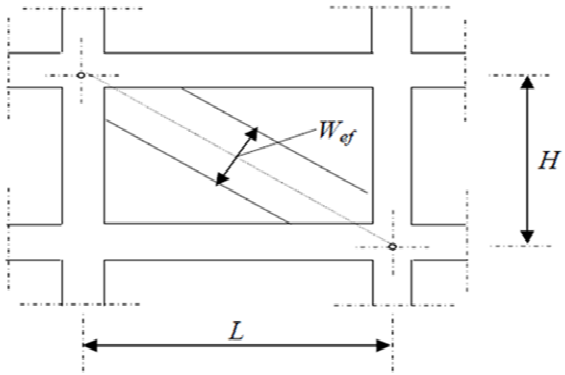


Fig. 2 Equivalent single strut

B. Multiple Strut Model:

It has been argued in the past that a single diagonal strut may not adequately simulate the complex interaction between frame and infill. A single diagonal connected between joints of frame along loaded diagonals, and bending moment released at both ends of the strut so that moment is not transferred to strut from the frame. The bending moments and shear forces in frame members cannot replicate using a single diagonal strut connecting the two loaded corners. Therefore, in order to obtain a physically more appropriate model, multi-strut models were proposed for carrying out dynamic analyses of infilled frames [Thiruvengadam].

The single strut model cannot predict the local failure in the frame members and underestimates the force resultants in the frame members while the multiple strut model predict the local failure in the frame members.

The infill is discretized into a grid of shear panels and each such panel, considered under the state of pure shear, is replaced by two diagonals, one acting in compression and the other in tension. When the division is done in such a way that the length and height of the wall are divided into a certain number of equal parts, then these diagonals form a set of multiple diagonal struts running in both directions and making equal angles to the horizontal. Both end of strut members are pin jointed and it carry only compressive forces. The areas of the struts are determined by the following equation given by Thiruvengadam.

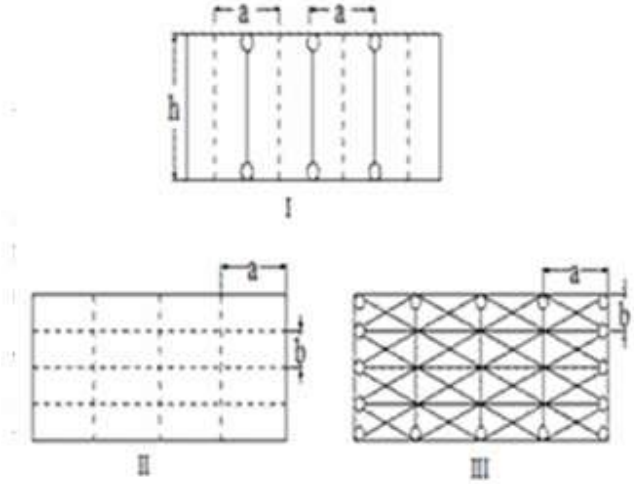


Fig. 3 Formulation of multiple strut

$$\text{Area of strut (Ad)} = \frac{a^2\, t}{4b\,(1+\mu)\cos^3\theta} \quad (1)$$

$$W = Ad/t \quad (2)$$

a = spacing of division of infill wall along length

b = spacing of division of infill wall along height

$$\theta = \tan^{-1}\left(\frac{b}{a}\right) \quad (3)$$

t = thickness of infill wall

μ= Poisson's ratio of infill

W= width of strut

III. SEISMIC ANALYSIS METHODS

There are different methods of analysis which provide different degree of accuracy. The analysis process can be categorized on the basis of three factors: the type of externally applied loads, the behaviour of structure/or structural materials and the type of structural model selected. Based on the type of external action and the behaviour of structure, the analysis can be further classified as linear static analysis, linear dynamic

analysis, nonlinear static analysis, or non-linear dynamic analysis.

Linear static analysis or equivalent static analysis can only be used for regular structure with limited height. Linear dynamic analysis can be performed in two way either mode superposition method or response spectrum method and elastic time history method. Non linear static analysis is an improvement over linear static analysis or dynamic analysis in the sense that it allows the inelastic behaviour of the structure. The method is relatively simple to implement and provides information on the strength, deformation and ductility of the structure and the distribution of demands. This permit to identify critical members likely to reach limit states during the earthquake, for which attention should be given during the design and detailing process. But this method contains many limited assumptions, which neglect the variations of loading patterns, the influence of higher nodes, and the effect of resonance. This method, under the name of push over analysis has acquired a great deal of popularity nowadays and in spite of these deficiencies this method provides reasonable estimation of the global deformation capacity, especially for the structures which primarily respond according to the first mode. A non-linear dynamic analysis or inelastic time history analysis is the only method to describe the actual behaviour of the structure during an earthquake.

IV. DESCRIPTION OF STRUCTURAL MODEL

For the purpose of this paper, a symmetrical reinforced concrete structure of a ten storey residential building is selected. The building is kept symmetric in both orthogonal directions in plan to avoid torsional response under pure lateral forces. The building considered being located in the seismic zone V and soil type is medium. For the purpose of this study the live load is taken to be 3 KN/m^2, the floor finish load is taken as 1.0 KN/m^2. Wind loading is not considered. The nonlinear equivalent static analysis is performed on different models. The unit weights for concrete and masonry are taken as 25 KN/m^3 and 18 KN/m^3 respectively. The elastic modulus of concrete is taken as 25000 MPa and that of masonry is taken as 4125 MPa. The Poisson's ratio for concrete is 0.2 and masonry is taken as 0.19. The total height of the building is 30 meter and height of each storey is 3 meters comprised of ten identical floors. The length of building is 20 meters while the width is 9 meters. The general layout is kept as regular as possible in order to focus an undistracted attention on the effect of the infill wall distribution. In the following study selected building models are analyzed by non-linear equivalent static method in SAP2000 software. In the following study five different models are numerically investigated; they vary in infill walls distribution.

Model I : Bare frame
Model II : Bare frame with full infill wall
Model III : Bare frame with soft storey
Model IV : Bare frame with full infill wall with Opening
Model V : Bare frame with soft storey with Opening

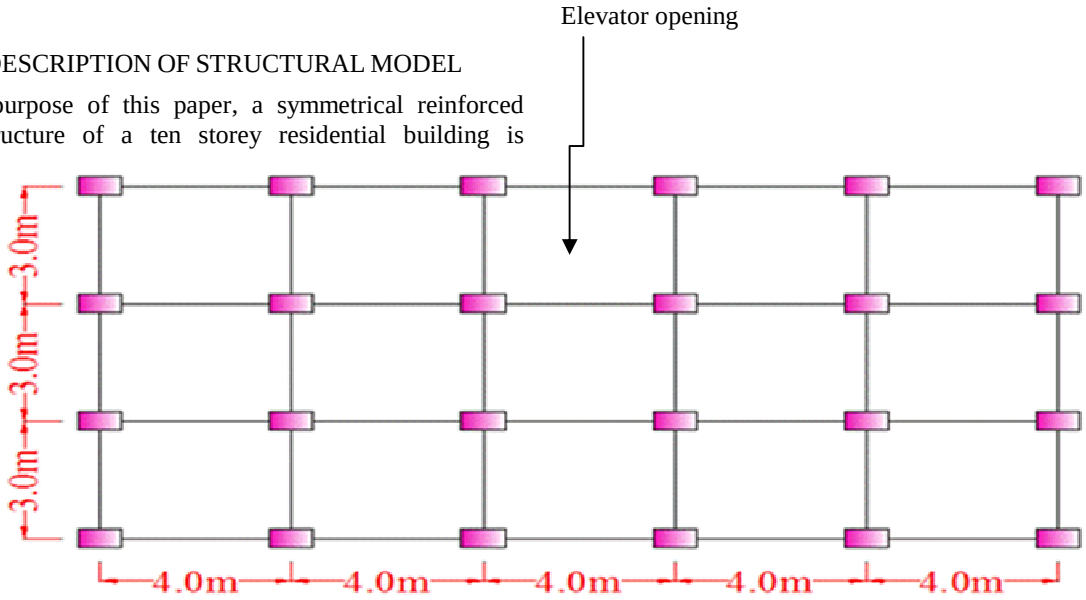


Fig.4 Floor plan of all structural model

In constructing various numerical models, columns having cross section of 700 x 300 mm (GL to 1st floor), 600 x 300 mm (2nd to 4th floor), 550 x 300 mm (5th to 7th floor) and 500 x 300 mm (8th to 9th floor) and all beams having cross section of 200 x 550 mm; slabs are modeled as shell element of 120 mm thickness. Walls are modeled as multiple strut of width 0.54 m for 4 m x 3 m panel and 0.44 m for 3 m x 3 m panel; thickness of wall is 0.23 m. Beams and columns are modeled as frame elements. The frame element is a two-noded (each having six degrees of freedom) element using a 3D beam column formulation which includes the effects of biaxial bending, torsion, axial deformation and biaxial shear deformations.

Slabs are modeled as shell elements having four-node formulation (each having four degrees of freedom). Walls are modeled as multiple strut having tension limit is zero; both end of strut is pin jointed. Window and doors openings are accounted for by removing or assigning zero areas for the struts crossing the opening area. The infill division into shear panels is to be done in such a way that the openings are covered by full panels.

V. RESULT AND DISCUSSION

A. Fundamental Natural Period

The stiffer structures have lesser natural period and their response is governed by the ground acceleration; most buildings fall in this category. The flexible structures have larger natural period and their response is governed by the ground displacement, for example, large span bridges. The approximate fundamental natural period of vibration of all buildings with brick infill panels may be estimated by the empirical expression from IS 1893-2002 (**clause -7.6.2**):

$$T_a = \frac{0.09 \times h}{\sqrt{d}}$$

Table I Comparison of fundamental period (Sec)

Mode	Bare	Full Infill	Soft Storey	Opening	Opening & Soft
1	1.08	0.45	0.48	0.542	0.56
2	0.855	0.27	0.355	0.369	0.439
3	0.835	0.225	0.295	0.364	0.378
4	0.294	0.115	0.141	0.152	0.164
5	0.278	0.09	0.105	0.125	0.145

B. Storey Drift

Storey drift is directly related to the stiffness of the structure. The higher the stiffness lowers the drift and higher the lateral loads on the structure. The value of the average storey drift of full infill frame is decrease by 73.75 % as compared to bare frame. There is a decrease of storey drift by 72.21 %, 54.15 % and 54.0 % in soft storey frame, frame with opening frame and frame with opening and soft storey frame than the bare frame.

C. Storey Stiffness

Stiffness is calculated by assuming that supports are fixed and load is applied at floor level. Horizontal displacement is measured at floor level and lateral stiffness is calculated by dividing horizontal deflection to lateral load.

Table II Storey drift due to EQx forces (mm)

Storey	Bare	Full Infill	Soft Storey	Opening	Opening & Soft
1	1.26	0.70	1.10	0.948	1.143
2	2.29	0.71	0.80	1.32	1.386
3	3.05	0.87	0.87	1.578	1.56
4	3.51	0.95	0.969	1.734	1.711
5	3.82	0.99	1.00	1.79	1.76
6	4.07	1.01	1.01	1.81	1.78
7	4.00	0.98	0.98	1.73	1.70
8	3.86	0.92	0.91	1.60	1.56
9	3.77	0.82	0.83	1.43	1.40
10	3.37	0.71	0.71	1.197	1.17

In other words stiffness is the force needed to cause unit displacement and is given by slope of force displacement relationship. The storey stiffness of full infill frame is generally higher as compared to other model. The averae storey stiffness of full infill frame increases by 90.78 % than bare frame. In soft storey frame average storey stiffness is increase by 81.19 % as compared to the bare frame. There is an increase of average storey stiffness by 29.59 % and 26.29 % in frame with opening and frame with opening & soft storey than bare frame. So that presence of opening in infill panel reduces the stiffness of frame.

Fig. 5 Storey stiffness (KN/mm)

D. Bending Moment and Axial Force in Column

Bending moment in column is governed by the load case such as 1.5DL+1.5EQX. The axial force in column is governed by 1.2DL+1.2EQX+1.2EQY. The result of the all load cases is given by considering the right corner column of the front bay of frame. The lateral forces in bare frame are transferred by frame actions and forces in infill frame are transferred by the truss action. So bending moment in bare frame is higher as compared to infill frame. The bending moments in columns also decrease due to presence of infill panel in the bare frame. There is a decrease of the average B.M in corner column by 28.47 % in full infill frame as compared to bare frame. In soft storey frame and frame with opening average bending moment decrease by 26.28 % and 8.27 % than bare frame. The bending moment decreases by 8.91 % in the opening & soft storey frame than bare frame. bare frame. The axial forces of stiff frame are higher as compared to bare frame. There is an increase of the average axial force of corner column by 8.78 % in full infill frame as compared to bare frame. In soft storey frame and frame with opening average axial force increases by 9.45 % and 10.36 % than the bare frame. As compared to bare frame axial force increases by 10.01 % in frame with opening & soft storey.

Table III Bending moment and axial forces in column

Storey	Bare		Full Infill		Soft Storey		Opening		Opening & Soft	
	Moment (KN.m)	Axial Force (KN)	Moment (KN.m)	Axial Force (KN)	Moment (KN.m)	Axial Force (KN)	Moment (KN.m)	Axial Force (KN)	Moment (KN.m)	Axial Force (KN)
1	145.62	1537.4	149.07	1920.5	150.73	1930.3	151.15	1832.8	144.6	1803.8
2	98.74	1402.2	77.68	1644.8	74.3	1680.7	96.9	1606.2	92.79	1632.2
3	95.1	1253.0	80.94	1407.4	83.73	1415.8	102.7	1421.6	102.3	1415.9
4	104.49	1095.3	80.34	1180.0	81.37	1182.2	107.2	1213.2	107.2	1207.5
5	107.95	927.83	76.84	959.68	77.05	959.4	105.1	1002.4	103.9	996.9
6	98.5	757.05	67.2	746.41	67.12	745.68	93.45	793.85	92.28	789.54
7	99.43	587.46	62.56	549.4	62.37	548.78	87.64	595.76	86.56	592.73
8	95.4	417.03	54.52	367.74	54.36	367.43	77.82	407.16	76.8	405.4
9	80.92	251.2	39.93	205.62	39.88	205.5	60.12	232.32	59.43	231.6
10	80.52	91.84	31.39	69.56	31.35	69.52	46.23	77.66	45.94	77.54

.

VI. CONCLUSIONS

- The results of present study show that the stiffness irregularity in infill panel increases the time period as compared to full infill panel.
- Presence of infill wall in frame decreased time period because of its higher stiffness as compared to bare frame. The stiffness of frame with opening and frame with opening & soft storey is very less as compared to the full infill frame.
- Storey drift of full infill frame decrease by 73.75 % than the bare frame. The storey drift of soft storey frame and frame with opening is higher due to less stiffness than the full infill frame. Storey drift of soft storey frame is increases due to the less stiffness at ground floor.
- The lateral forces in bare frame are transferred by frame actions and the forces in infill frame are transferred by truss action. The bending moment in corner column decrease by 28.47 % in full infill frame than bare frame. The opening and soft storey presents in infill panel increases bending moment in column than full infill frame.
- The axial forces of stiff frame are higher as compared to bare frame. There is an increase of the average axial force of corner column by 8.78 % in full infill frame as compared to bare frame.
- Thus modeling of a building considering infill shows improved performance of building during earthquake. The care shall be taken while modeling infill, as position of opening or presence of soft storey considerably changes the behaviour of structure.

VII. REFERENCES

[1] J.N. Arlekar, S.K. Jain and C.V.R. Murthy, “Seismic response of RC frame buildings with soft first storeys”, Proceedings of the CBR golden jubilee conference on natural hazards in urban habitat, 10-11 November, New Delhi,1997

[2] P. Agarwal and M. Shrinkhade, “Earthquake resistant design of structures“, Eastern economy edition, PHI press, New Delhi,2007

[3] A Habibullah. SAP 2000 "Static and dynamic finite element analysis of structures", Computers and structures inc. Berkeley, California,2005

[4] Samir Helou and A.R. Touqan, “Dynamic behaviour of reinforced concrete structures with masonry walls”, An- Najah university j. res. (N.Sc.), Vol. 22, pp. 78-92,2008

[5] Kasım, Fuat Demir, “Earthquake assessment of r/c structures with masonry infill walls”, International journal of science & technology, Vol. 2 (2), pp.155-164,2007

[6] V. Thiruvengadam, “The natural frequency of infilled frames”, Proceeding of earthquake engineering and structural dynamics, Vol. 13, pp. 401-419,1984

[7] A.L. Wilson, “Three-dimensional static and dynamic analysis of structures”, Computers and structures inc. Berkeley, California, 2002

CONSERVATION OF ENERGY BY LED LIGHTING & ITS ADVANTAGES

Prepared by Sagar J. Kothari (Lecturer Electrical Department, G.I.T.)

INTRODUCTION

In the time of Globalization, Energy Crises is hammering issue for us. In such time, use of LED light can serve as one of the alternate to conserve the Energy. Here one of such concept is mentioned.

We found that, there are large numbers of Street Lights in our city or town or now in our streets also. Mainly we use Tube light or Sodium Lamp as the Source of Street Light. But the life span of such light is very small. More they consume high amount of Electricity. And the main draw back is that they are not ECO friendly at the end or mean time.

There is millions of Street Light installed in cities only. So, you can think of your State and Our Country too. Therefore some basic, calculations, savings and advantages are mentioned in this report. Because this is the sector on which the countries like US, Canada, UK, Australia, South Africa, Nigeria and many other have started working and they are also replacing there old traditional lights to LED Lights.

Benefits of LED Lighting

Long life times and slow failure: LEDs have a long useful life of up-to 50,000-100,000 hours (though time to complete failure may be longer). CFL bulbs and tubes are rated at about 15,000 hours and incandescent light bulbs at about 1,500 hours. Also, LEDs fail by dimming over time, rather than the abrupt burn-out of incandescent bulbs.

Reduces Toxic Waste: In addition to the saved replacement costs, LEDs save on the disposal of incandescent bulbs and mercury tainted, compact fluorescent bulbs thus cleaning up our landfills and saving our environment from toxic waste. Also, CFLs and incandescent bulbs can become a hazard when broken.

Luminous efficacy: LEDs produce significantly more light (measured in lumens) per watt (140 lm/W & growing) than CFL (54 lm/W) and incandescent bulbs (11 lm/W).

Power efficiency: LEDs use significantly less power (watts) to produce the same amount of light (lumens), LED lights use 85% less energy than incandescent bulbs and a third of the energy of compact fluorescent lamps. Consumption of less energy results in fewer greenhouse gas emissions being released into our environment.

Wide temperature range: LEDs have a very wide temperature range (-30 to +50° C) and for that reason is the only source of light for deep freeze containers.

Versatile color: LEDs can emit light of an intended color without the use of color filters that traditional lighting methods require. This is more efficient and can lower initial costs.

Size and design flexibility: LEDs are very small (smaller than 2 mm) and provide for improved design flexibility as they are easily populated onto printed circuit boards.

Rapid On/Off time: LEDs have rapid on and off times achieving full brightness in nanoseconds.

Cycling: LEDs are the ideal one for the applications which are subjected to frequent on-off cycling. Unlike fluorescent lamps those burn out more quickly when cycling is frequent. More HID lamps require a long time before restarting.

Dimming: LEDs can very easily be dimmed with digital control.

Cool light: In contrast to most light sources, LEDs radiate very little heat that can cause damage to sensitive objects or fabrics. Energy is dispersed as heat through the base of the incandescent bulbs wasting 98% of their energy.

No radiation: LEDs do not emit infrared or UV in beam output unlike traditional forms and, as a consequence, they do not attract insects and other bugs…!!!!

Robust to shock: LEDs, being solid state components, are difficult to damage with external thermal and vibration shock, unlike fluorescent and incandescent bulbs which are fragile.

Focus: The solid package of the LED can be designed to focus its light. Incandescent and fluorescent sources require an external reflector to direct it in a usable manner. Directed light output in LEDs provide for increased system efficiency.

Pure non-flickering light: LEDs always get a pure, clear and consistent light that makes reading or viewing objects easier on the eyes. No annoying flickering caused by compact

.

fluorescent bulbs - a problem that has been blamed for migraine headaches and epileptic fits.

Toxicity: LEDs do not contain mercury or lead, unlike other lighting sources.

Green friendly: LEDs are ideal for use with alternative or renewable energy resources like solar and wind power.

COMPARISON ANALYSIS BETWEEN LED STREET LIGHT

AND CONVENTIONAL STREET LIGHT

ITEM	HIGH PRESSURE SODIUM LAMP	LED STREET LIGHT
Photometric Performance	Bad	Excellent
Radiator Performance	Bad	Excellent
Electrical Performance	Electric Shock Easy (High Voltage)	Safe (Low Voltage)
Working Life	Short (5000 hrs)	Very Long (> 50000 Hrs)
Working Voltage Range	Narrow (+/- 7%)	Wide (+/- 20%)
Power Consumption	Quite Height	Quite Low
Startup Speed	Quite Slow (Over 10 Minutes)	Rapid (2 Second)
Strobe	Yes (Alternating Current Drive)	No (Direct Current Drive)
Optical Efficiency	Low (<60%)	High (>90%)
Color Index / Distinguish Feature	Bad, Ra<50 (The Color of Object is Faith, Boring, Hypnosis)	Good, Ra>75 (The Color of object Is Fresh, Veritable and Comfortable)
Color Temperature	Quite Low (Yellow or Amber, Uncomfortable)	Ideal Color Temperature (Comfortable)
Bad Glare	Strong Glare	No Glare
Light Pollution	Serious	NO
Heat Generation	Serious (>300' C)	Cold Light Source (<60' C)
Lampshade Turn Dark	Easy (Dust Absorption)	No (Static Proof)
Lampshade Aging Turn Yellow	Very Fast	Not
Shockproof Performance	Bad (Frangibility)	Good (No Filament Nor Glass)
Environment Pollution	Lead Pollution, Etc.	None
Maintanance Costs	High	Quite Low
Product Cubage	Very Large	Small (Slim Appearance)
Product Weight	Heavy	Light
Cost Effective	Bad	High
Integrated Performance	Bad	Excellent

LED BASED STREET LIGHT

LED BASED FLOOD LIGHT

The Life span of the LED is minimum 50,000 hours (i.e. 50,000hours/365days=136.98years, lets say daily working hours are 10, so 136.98 years/10hours=13.69Years. Even though we have calculated for 10 Years)

Energy Consumption Calculation:-

List	Sodium Vapor Lamp	LED Lamp	Saving
Rating	70W	24W	46W
Net Consumption	85W	28.8W	56.2W per Hour per One Bulb
Total Hours of	10 Hours	10 Hours	

Operation/day			
Daily Consumption	85W × 10hr = 0.85 KWhr = 0.85 Units	28.8 × 10hr = 0.28 KWhr = 0.28 Units	0.57 Units are saved n daily basis.
Per Unit Prize	Rs. 4.50	Rs. 4.50	
Daily Bill	Rs. 3.83	Rs. 1.26	Rs. 2.57
Monthly Bill	Rs. 114.9	Rs. 37.80	Rs. 77.10
Yearly Bill	Rs. 1,378.80	Rs. 453.6	Rs. 925.20
Life	Max. 1 Year	50,000 hours = Min 10 Years	
Prize	Rs. 1,500/-	Rs. 15,000/-	
1st Year Expense	Rs. 2878.8 (Bulb Prize 1500+ Bill 1378.80)	Rs. 15,453.60 (Bulb Prize 14000+ Bill 453.60)	Rs. -12,574.80
2nd Year Expense	Rs. 5757.60 (Last Year 2878.8 + 2878.8 of this year also)	Rs. 15,907.20 (15453.60 of last year + 453.60 of this year bill)	Rs. -10,149.60
3rd Year Expense	Rs. 8,636.40	Rs. 16,360.80	Rs. -7,724.40
4th Year Expense	Rs. 11,515.20	Rs. 16,814.40	Rs. -5,299.20
5th Year Expense	Rs. 14,394.00	Rs. 17,268.00	Rs. -2,874.00
6th Year Expense	Rs. 17,272.80	Rs. 17,721.60	Rs. -448.80
7th Year Expense	Rs. 20,151.60	Rs. 18,175.20	Rs. 1,976.40
8th Year Expense	Rs. 23,030.40	Rs. 18,628.80	Rs. 4,401.60
9th Year Expense	Rs. 25,909.20	Rs. 19,082.40	Rs. 6,826.80
10th Year Expense	Rs. 28,788.00	Rs. 19.536.00	Rs. 9,252.00

Energy Consumption Calculation:-

List	Sodium Vapor Lamp	LED Lamp	Saving
Rating	150W	35W	115W
Net Consumption	180W	42W	138W per Hour per One Bulb
Total Hours of Operation/day	10 Hours	10 Hours	
Daily Consumption	180W × 10hr = 1.8 KWhr	42 × 10hr = 0.42 KWhr	1.38 Units are saved on daily basis.
Per Unit Prize	Rs. 4.50	Rs. 4.50	
Daily Bill	Rs. 8.10	Rs. 1.89	Rs. 6.21
Monthly Bill	Rs. 243.00	Rs. 56.70	Rs. 186.30
Yearly Bill	Rs. 2,916.00	Rs. 680.40	Rs. 2,235.60
Life	Max. 1 Year	50,000 hours = Min 10 Years	
Prize	Rs. 2,100/-	Rs. 18,700/-	
1st Year Expense	Rs. 5,016 (Bulb Prize 2100+ Bill 2916)	Rs. 19,380.40 (Bulb Prize 18700+ Bill 680.40)	Rs. -14,364.40
2nd Year Expense	Rs. 10,032 (Last Year 5016 + 5016 of this year also)	Rs. 20,060.80 (19380.40 of last year + 680.40 of this year bill)	Rs. -10,028.80
3rd Year Expense	Rs. 15,048.00	Rs. 20,741.20	Rs. -5,693.20
4th Year Expense	Rs. 20,064.00	Rs. 21,421.60	Rs. -1,357.60
5th Year Expense	Rs. 25,080.00	Rs. 22,102.00	Rs. 2,978.00
6th Year Expense	Rs. 30,096.00	Rs. 22,782.40	Rs. 7,313.60
7th Year Expense	Rs. 35,112.00	Rs. 23,462.80	Rs. 11,649.20
8th Year Expense	Rs. 40,128.00	Rs. 24,143.20	Rs. 15,984.80
9th Year Expense	Rs. 45,144.00	Rs. 24,823.60	Rs. 20,320.40
10th Year Expense	Rs. 50,160.00	Rs. 25,504.00	Rs. 24,656.00

Energy Consumption Calculation:-

List	Sodium Vapor Lamp	LED Lamp	Saving
Rating	250W	48W	202W
Net Consumption	300W	57.6W	242.6W per Hour per One Bulb
Total Hours of Operation/day	10 Hours	10 Hours	
Daily Consumption	300W × 10hr = 3.0 KWhr = 3.0 Units	57.6 × 10hr = 0.576 KWhr = 0.58 Units	2.42 Units are saved on daily basis.
Per Unit Prize	Rs. 4.50	Rs. 4.50	

Daily Bill	Rs. 13.50	Rs. 2.61	Rs. 10.89
Monthly Bill	Rs. 405.00	Rs. 78.30	Rs. 326.70
Yearly Bill	Rs. 4,860.00	Rs. 939.60	Rs. 3,920.40
Life	Max. 1 Year	50,000 hours = Min 10 Years	
Prize	Rs. 2,600/-	Rs. 22,800/-	
1st Year Expense	Rs. 7,460 (Bulb Prize 2600+ Bill 4860)	Rs. 23,739.60 (Bulb Prize 22800+ Bill 939.60)	Rs. - 16,279.60
2nd Year Expense	Rs. 14,920 (Last Year 7460 + 7460 of this year also)	Rs. 24,679.20 (last 23739.60 year + 939.60 bill)	Rs. - 9,759.20
3rd Year Expense	Rs. 22,380.00	Rs. 25,618.80	Rs. - 3,238.80
4th Year Expense	Rs. 29,840.00	Rs. 26,558.40	Rs. 3,281.60
5th Year Expense	Rs. 37,300.00	Rs. 27,498.00	Rs. 9,802.00
6th Year Expense	Rs. 44,760.00	Rs. 28,437.60	Rs. 16,322.40
7th Year Expense	Rs. 52,220.00	Rs. 29,377.20	Rs. 22,842.80
8th Year Expense	Rs. 59,680.00	Rs. 30,316.80	Rs. 29,363.20
9th Year Expense	Rs. 67,140.00	Rs. 31,256.40	Rs. 35,883.60
10th Year Expense	Rs. 74,600.00	Rs. 32,196.00	Rs. 42,404.00

Daily Bill	Rs. 7.86	Rs. 1.35	Rs. 6.51
Monthly Bill	Rs. 235.80	Rs. 40.5	Rs. 195.30
Yearly Bill	Rs. 2839.60	Rs. 486.00	Rs. 2,353.60
Life	Max. 1 Year	50,000 hours = Min 10 Years	
Prize	Rs. 5,000/-	Rs. 10,000/-	
1st Year Expense	Rs. 7,839.60 (Bulb & Fixture Prize 5000+ Bill 2839.60)	Rs. 10,486.00 (Bulb Prize 10000+ Bill 486)	Rs. - 2,646.40
2nd Year Expense	Rs. 11,679.2 (Last Year 7839.60 + 1000 of Bulb + Bill 2839.60)	Rs. 10,972.00 (last 10,486 year + 486 bill)	Rs. 707.20
3rd Year Expense	Rs. 15518.80	Rs. 11,458.00	Rs. 4,060.80
4th Year Expense	Rs. 19,358.40	Rs. 11,944.00	Rs. 7,414.40
5th Year Expense	Rs. 23,198.00	Rs. 12,430.00	Rs. 10,768.00
6th Year Expense	Rs. 27,037.60	Rs. 12,916.00	Rs. 14,121.60
7th Year Expense	Rs. 30,877.20	Rs. 13,402.00	Rs. 17,475.20
8th Year Expense	Rs. 34,716.80	Rs. 13,888.00	Rs. 20,828.80
9th Year Expense	Rs. 38,556.40	Rs. 14,374.00	Rs. 24,182.40
10th Year Expense	Rs. 42,396.00	Rs. 14,860.00	Rs. 27,536.00

Energy Consumption Calculation for Flood Light:-

List	**Metal Halide Lamp**	**LED Lamp**	**Saving**
Rating	150W	24W	126W
Net Consumption	175W	30W	145W per Hour per One Light
Total Hours of Operation/day	10 Hours	10 Hours	
Daily Consumption	175W × 10hr = 1.75 KWhr = 1.75 Units	30 × 10hr = 0.3 KWhr = 0.3 Units	1.45 Units are saved on daily basis.
Per Unit Prize	Rs. 4.50	Rs. 4.50	

Conclusion:-

Thus we can conclude that the LED based Lights can save three times more power than traditional Light and also produce three times less bill as compared to the other one. While Life of the LED lights and other mentioned advantages prove its characteristics. More, unlike a bulb if it gets fail it doesn't give illumination while if one or couple of LEDs get fail in the LED Light it doesn't effect on the other LEDs.

Gandhinagar Institute of Technology

We Always Believe in Quality Education

collage

Contact US

Address:

Gandhinagar Institute of Technology
Khatraj - Kalol Road,
Village Moti Bhoyan,
Tal. Kalol
Gandhinagar-382721

Telephone No :
+91-02764-281860
+91-02764-281861

Web Site: **www.git.org.in**
E-mail: jet@git.org.in

Fax:
+91-02764-281862