



PLATINUM FOUNDATION MANAGED

# GANDHINAGAR INSTITUTE OF TECHNOLOGY

GIT-JOURNAL OF ENGINEERING AND TECHNOLOGY (2012 – FIFTH VOLUME)

NATIONAL LEVEL RESEARCH JOURNAL – ISSN 2249-6157

[Home](#)

[Trustee](#)

[Editorial Board](#)

[Director Message](#)

[Papers](#)

[Contact Us](#)



## About Gandhinagar Institute of Technology

Gandhinagar Institute of Technology is established by Platinum Foundation in 2006. It offers under graduate programs in Mechanical Engineering, Information Technology, Computer Engineering, Electronics and Communication Engineering and Civil Engineering and Post graduate program in MBA (Finance, Human Resource Development, and Marketing) and M.E. in Mechanical Engineering with specialization in Thermal Engineering and Computer Aided Design & Computer Aided Manufacturing.

All these programs are approved by AICTE, New Delhi and affiliated to Gujarat Technological University. We have elaborate laboratory facilities and highly motivated and qualified faculty members. We are also arranging technical seminars, conferences, industrial institute interaction programs, workshops and expert lectures of eminent dignitaries from different industries and various reputed educational institutes.

Our students are innovative and have excellent acceptability to latest trends and technologies of present time. Our students have also participated in various technical activities as well as sports activities and have achieved various prizes at state level. We have two publications a National level research journal 'GIT-Journal of Engineering and Technology (ISSN 2249-6157)' and 'GIT-a Song of Technocrat' (college magazine).



PLATINUM FOUNDATION MANAGED

# GANDHINAGAR INSTITUTE OF TECHNOLOGY

GIT-JOURNAL OF ENGINEERING AND TECHNOLOGY (2012 – FIFTH VOLUME)

NATIONAL LEVEL RESEARCH JOURNAL – ISSN 2249-6157

<a href="#">Home</a>	<a href="#">Trustee</a>	<a href="#">Editorial Board</a>	<a href="#">Director Message</a>	<a href="#">Papers</a>	<a href="#">Contact Us</a>
----------------------	-------------------------	---------------------------------	----------------------------------	------------------------	----------------------------

## Shri Harishbhai B. Rohera, Trustee

**Qualifications** : B. Com

### Background

- » Proprietor: Mahadev Steel Suppliers
- » C/o: Vinayak Steel Syndicate
- » C/o: Dhiraj Steel Supplier
- » C/o: Krishna Steel Trader
- » Owner: National Steel Processor
- » Trustee: Sai Vasant Ghot Darbar
- » Trustee: Jai Jhulelal Mandir



## Shri Gahanshyambhai V. Thakkar, Trustee

**Qualifications** : M.A

### Background

- » Professor at Vivekanand College of Arts, Ahmedabad
- » Ex. M.L.A., Gujarat Assembly from Mandal
- » Trustee of V.M. Thakkar Charitable Trust, Ahmedabad
- » Manages Muktajeevan Vidhyalaya and BVD High School, Isanpur and Maninagar
- » Advisor/Member Kankaria Maninagar Nagarik Sahakari Bank
- » Director - Adarsh Co-Operative Departmental Stores



## Shri Dipakbhai N. Ravani, Trustee

**Qualifications** : B.Com., LL.B.

### Background

- » Business



## Shri Pravinbhai A. Shah, Trustee

**Qualifications** : B.A., LL.B.

### Background

- » President of Zalavad Samaj Jain Seva Trust
- » Trustee of Vasant » Atma Charitable Trust
- » Trustee of Rampura Champa Vijaya Hospital
- » Trustee of Shantivan & Ambawadi Jain Sangh
- » Trustee of Rampura Kelavani Mandal
- » Trustee of Pampura Panjrapole Trust



## Smt Varshaben M. Pandhi, Trustee

**Qualifications** : B.Com

### Background

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



## Mr. Mahendrabhai Pandhi, Member of Governing Body

**Qualifications** : B.Com, F.C.A

### Background

- » Proprietor, M. R. Pandhi & Associates
- » He has many Indian clients having international presence
- » His areas of interest are Taxation, Audit, Project Finance and Company Law related matters.
- » He is one of the members of the study group of 25 Chartered Accountants constituted by WIRC
- » He has visited many countries like U.A.E., Moratus, Singapore and Africa for his client work.





PLATINUM FOUNDATION MANAGED

# GANDHINAGAR INSTITUTE OF TECHNOLOGY

GIT-JOURNAL OF ENGINEERING AND TECHNOLOGY (2012 – FIFTH VOLUME)

NATIONAL LEVEL RESEARCH JOURNAL – ISSN 2249-6157

<a href="#">Home</a>	<a href="#">Trustee</a>	<a href="#">Editorial Board</a>	<a href="#">Director Message</a>	<a href="#">Papers</a>	<a href="#">Contact Us</a>
----------------------	-------------------------	---------------------------------	----------------------------------	------------------------	----------------------------

## Editorial Board

Prof. M. N. Patel Chairman	Principal L.D. College Of Engineering Navrangpura, Ahmedabad - 380 015.
Prof. Paritosh K. Banik	Director General Pandit Deendayal Petroleum University Raisan Village, Gandhinagar - 382 007.
Dr. Ketan Kotecha	Director Institute Of Technolgy, Nirma University Sarkhej – Gandhinagar High Way, Ahmedabad - 382481.
Dr. F. S. Umarigar	Principal Birla Vishvakarma Mahavidhyalaya Vallabhvidyanagar, Anand - 388120.
Dr. N M Bhatt Editorial Chief	Director Gandhinagar Institute of Technology Moti Bhoyan, Ta. Kalol Gandhinagar - 382721.
Disclaimer Views expressed in the papers are solely from respective authors and editorial board has no responsibility of their authentication This journal is being published in the interest of research community for private circulation only	



PLATINUM FOUNDATION MANAGED

## GANDHINAGAR INSTITUTE OF TECHNOLOGY

GIT-JOURNAL OF ENGINEERING AND TECHNOLOGY (2012 – FIFTH VOLUME)

NATIONAL LEVEL RESEARCH JOURNAL – ISSN 2249-6157

[Home](#)

[Trustee](#)

[Editorial Board](#)

[Director Message](#)

[Papers](#)

[Contact Us](#)

### Message from Director



It gives me immense pleasure that the fifth issue of our National journal 'GIT-Journal of Engineering and Technology' is being published with ISSN 2249 – 6157 for fifth successive year. The annual journal contains peer reviewed technical papers submitted by the researcher of all domains of engineering and technology. The issue is a result of imaginative and expressive skill and talent of GIT family. Research papers were invited from the researcher of all domains of engineering and technology. More than 190 research papers were received. After peer review about 82 papers are selected and are being published in this issue of the journal.

GIT was established in 2006 and during a short span of six years; it has accomplished the mission effectively for which it was established. Institute has been constantly achieving the glory of excellence in the field of curricular, co-curricular and extra-curricular activities. For the fourth consecutive year an annual technical symposium TechXtreme-2012 was successfully organized by the institute. More than 1200 students of various technical institutions across the Gujarat participated in the TechFest. Prizes worth Rs 2 lacs and trophies were given to the winners of total 36 events. During the year institute has organized Spoken tutorial on Linux, Latex, Scilab, and Python in association with IIT Bombay, seminar on Cisco Networking by Network Nuts, Robotics workshop in association with Star Robotics, CAD/CAM workshop in association with Khodiyar CAD Center for its students. The institute has also successfully organized Debate Competition, Rangoli Competition, Kite Flying competition, Ratri B4 Navaratri, and Sports activities. Institute has also arranged two blood donation drives and more than 300 units were collected from the students and staff members. Students have also participated and won prizes in various sports event organized by other Institutions including that of GTU. Students of the institutes won prizes in many technical symposiums organized at various engineering colleges of Gujarat. Institute has organized many industrial visits and expert lectures for the students for supplementing the class room teaching. I am extremely happy to mention that throughout the year the faculty members have worked very hard to achieve all kinds of curricular, co-curricular and extra-curricular activities.

The Institute is also emphasis on academic development of its faculty members. During the year, **12** International and **15** National papers were presented by the faculty members at various conferences organized across the India. The faculty members were also deputed to attend total **120** seminars/workshops/training programs/symposiums. The institute has organized many state level seminars and workshops on current trends of Engineering and Management. Spoken tutorial on Linux, Latex, Scilab, and Python in association with IIT Bombay, CAD/CAM/CAE workshop by Khodiyar CAD Center and AutoCAD 2011 Professional Certificate examination by Auto Desk are few of them.

Publication of the journal of national level is not possible without whole hearted support of committed and experienced Trustees of Platinum Foundation Mr. Hareshbhai Rohera, Mr. Ghanshyambhai Thakkar, Mr. Deepakbhai Ravani, Mr. Pravinbhai Shah and Smt. Varshaben M. Pandhi. I take an opportunity to express my deep feelings of gratitude to all the trustees of Platinum Foundation and Mr. Mahendrabhai Pandhi, member of Governing body of the trust for their constant support and motivation.

It's my privileged to compliment the staff members and the students for showing high level of liveliness throughout the year. I also congratulate the team of the 'GIT-Journal of Engineering and Technology' for their untiring effort to bring out this fifth issue of the journal.

Dr N M Bhatt  
Director

# STUDY OF GEL TIME AND STRENGTH FOR POLYMER BASED SODIUM SILICATE GROUT AND COLLOIDAL SILICA GROUT USED IN SANDY SOIL

A. Komal.k.Dave

Asst.professor, civil department  
Babaria Institute of Technology,  
BIT campus  
Vadodara-Mumbai NH#8  
Vadodara, India  
xpkomal@yahoo.in

B.Hiral.S.Modha

Asst.professor, civil department  
Babaria Institute of Technology  
BIT campus  
Vadodara-Mumbai NH#8  
Vadodara, India  
Modhahiral8@gmail.com

c.Nipa.A.Desai

Asst.professor, civil department  
Babaria Institute of Technology  
BIT campus  
Vadodara-Mumbai NH#8  
Vadodara, India

**Abstract**—Chemical grouting is a soil stabilization technique to fill voids in the ground with aim to increase resistance against deformation, to supply cohesion, shear strength and uniaxial compressive strength. Chemical grouting is the type of grouting, in which chemicals are used. Sodium silicate is a popular non toxic and non corrosive chemical grout and free from health and environmental effects, while Colloidal silica is an innovative chemical grout material. Chemical grout are used for stabilizing the load bearing capacity of fine grained materials in foundation and the control of water in mine shafts, tunnels, trenches, and other excavation. Various lab experiments and studies are performed to check properties of different grout materials and their actual performance. The paper is focusing on the results obtained from study and lab experiments conducted for Polymer based sodium silicate and colloidal silica with natural salt as reactants i.e. Cao, for sandy soil. The tests i.e. gel time, specific gravity, pH test, syneresis, UCS for raw gel and grouted sand are performed. From the result obtained gel time and UCS of both the grout mix are studied and comparison for gel time and strength is presented.

**Key words:** Colloidal silica, Sodium silicate, Gel time, UCS (unconfined compressive strength)

## I. INTRODUCTION

Grouting is one of the improvement techniques, carried out in relatively permeable and weak foundation. Chemical grout were developed in response to a need to develop strength and control water flow in geologic units where the pore size in the rock or soil units were too small for conventional Portland cement suspensions. Chemical grouting is principally used for rehabilitation work under and around damaged or deteriorated structures.

## A. Introduction of material used

The sodium silicate and colloidal silica are used as base material; their properties are given in table.1 and table.3

Table no.1 Properties of sodium silicate.

Constituents	Percentage
Specific gravity	Be59
Na <sub>2</sub> (% wt.)	22.17
SiO <sub>2</sub> (% wt.)	35.26
Na <sub>2</sub> O :SiO <sub>2</sub>	1:1.59
Moral ratio	1:1.64
Total solids	57.43
Density g/cc	1.75

In present experiment work, Bahadurpur sand is used for all the experiments. The sand passing through 425 $\mu$  and retained on 75 $\mu$  is used for preparing grouted sand samples.

Table no.2 Properties of fine sand.

Properties	Fine sand
Specific gravity	2.59
Coefficient of curvature(Cc)	0.72
Uniformity coefficient(Cu)	2.17

Table no.3 properties of colloidal silica.

Test Grade	CILICOL 30 AK
Concentration of SiO <sub>2</sub> % wt	30-31
Concentration Na <sub>2</sub> O % wt	0.3-0.5
pH	9.5-10.5
Particle Size (nm)	10-20
Viscosity (cps)at 25E	<5
Sp. Gravity at 20	1.20-1.22
Appearance	Clear to Opalescent
Stability	Semi Permanent

For experimental work, dose of chemical admixture were varied and dose of sodium silicate was kept constant and four grout mixes were prepared. Different ratio of water and colloidal silica grout was taken by changing the ratio. Grouted sand samples were prepared by injecting grout through syringe in the fine sand.

### B. Introduction of grouting material

Sodium silicate comes in a form of syrupy liquid and its viscosity varies with water content and ratio of SiO<sub>2</sub>-Na<sub>2</sub>O, silicate is a weak acid and consequently sodium silicate is a basic. Silicate will be precipitated in form of a virtually firm gel by neutralizing with strong and weak acids. Gellification is also obtained by the addition of bivalent or trivalent cations. It is necessary to delay the setting time of gel, and this is accomplished by diluting the silicate and reducing the quantity of reactive material added to it. This produces a soft gel which is useful only for reducing the permeability of the soil. The Colloidal Silica (CS), which is a silicon-based chemical grout. It poses no health hazard, is unaffected by filtration is chemical and biologically inert, has excellent durability characteristics. Colloidal silica is an innovative grout. This is made by extracting alkali from sodium silicate using ion- exchange resin, in the factory. This colloidal silica has an electrical double layer around its surface. After breaking the layers with inorganic salt the colloidal are bonded with siloxane bonds and thus develop a gel network. The size of colloidal silica is about 10 to 100 nm in diameter. The network of this colloidal silica grout is formed by condensation and polymerization of silanol radicals on the surface of the colloidal. Consequently, the structure of colloidal silica grout is considered to be a pile of spherical connected to each other.

## II. OBJECTIVE OF STUDY

The present investigation has been carried out to study gel time, syneresis, pH, time- viscosity and Unconfined Compressive Strength for CS grout using Cao as reactant among all reactants( Calcium Chloride, Magnesium Chloride, Potassium Chloride, Calcium Oxide, and ordinary Portland cement) and to design an economical, non toxic chemical grout using sodium silicate as base and polymaleic acid (polymer) as a catalyst, mineral acid as an additive, and calcium chloride as a hardner.The experiment work include finding out various physical properties like gel time, Ucs for raw gel and grouted mass. It includes finding the relation between the gel time and strength of raw grout and grouted mass.

## III. METHDOLOGY

The proportion of the both grout materials are taken as shown in table No.4 and table no.5 for the analysis.

Grout type (100 ml)	Sodium silicate (ml)	Water (ml)	Acrylic acid (ml)	Mineral acid (ml)	Calcium chloride (gm)
Grout-A	25	75	2	2	2
Grout-B	25	75	3	3	2.5
Grout-C	25	75	4	4	1.5
Grout-D	25	75	5	5	0.7

Table no.4 Proportion of chemicals and admixture used for analysis for polymer based sodium silicate grout.

Grout type	Colloidal silica (ml)	Water (ml)	W/ CS	Calcium oxide (gm)
Grout-A	100	50	0.5	2.5
Grout-B	50	50	1.0	2.5
Grout-C	50	100	2.0	2.5

Table no.5 Proportion of chemicals and admixture used for analysis for colloidal silica.

The time of gelation is measured by the conventional beaker pouring method as standardized by Barbadette (1955) by the deformation of 100 cc grout in a cup of 50 mm diameter placed in three positions - horizontal, inclined and inverted to access the time for partial and full gelation. This method is found suitable for thick grouts. In any grout the gel time can be varied to some extent by varying its ingredients like precipitant, activator, accelerator, and retarder. The conventional method of obtaining cylindrical samples for unconfined compressive strength test cannot be used to prepare grout mix. For this test PVC pipes were used to prepare the cylinder raw grout samples. The internal diameter of the PVC pipe was 38 mm. After pouring the grout into PVC pipe, it was kept vertical under wet cured condition

for different curing period, i.e. 3,7,14 and 28 days for polymer based sodium silicate grout and 3,7,30 and 90 days for CS grout. The sample is extracted from the pipe and is trimmed to get horizontal surface and to make it exact 76 mm high so as to maintain  $H/D = 2$ . The care is taken to keep the top and the bottom surfaces exactly perpendicular to the axis of the specimen. The length of the sample is kept more than that required which is shown in figure-2. All grout samples are tested on strain controlled triaxial compression machine at strain rate of 0.04 mm/ min., after mentioned curing period. Two or three samples are tested for each curing time. Fig.1 and Fig.2 show the gel time test and raw grout sample respectively.



Fig.1 Gel time test Fig.2 Raw grout sample

**IV. ANALYSIS**

*A. Polymer based sodium silicate grout*

Fig:3 shows the effect of Polymaleic acid and mineral acid on gel time of the grouts. The concentration of calcium chloride is different for different grout type.

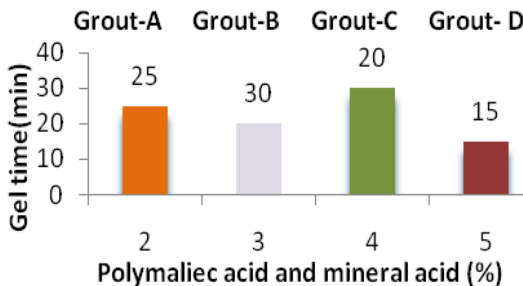


Fig.3 geltime v/s concentration of polymaleic acid and mineral acid

Unconfined compression tests were performed to determine the strength of raw grouts. Fig.5 shows the Peak strength versus curing time for unconfined compression test after 3, 7, 14 and 28 days for raw grout. Unconfined compression test were performed for all types dry grouted sand at 3,7,14 and 28 days.

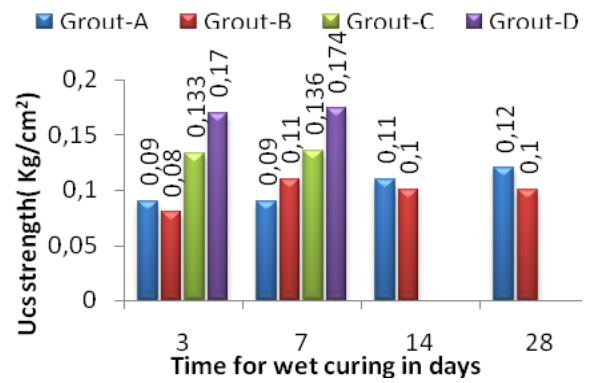


Fig: 4 UCS strength v/s time for raw grout

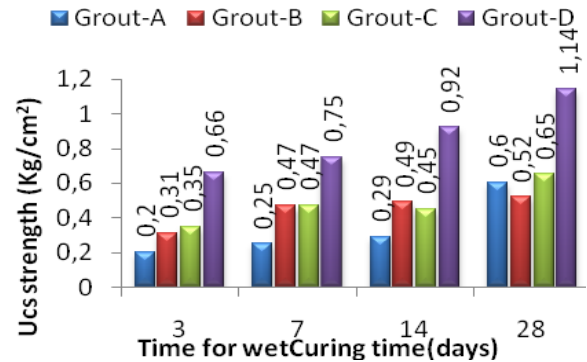


Fig: 5 UCS strength v/s time for grouted sand

*B. Colloidal silica grout*

Fig: 4 show the effect of concentration of CS on gel time with Cao as reactant.

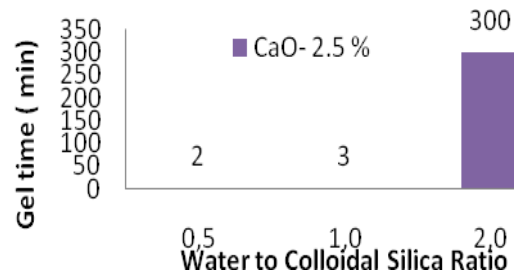


Fig: 6 Gel time v/s concentration of cs grout

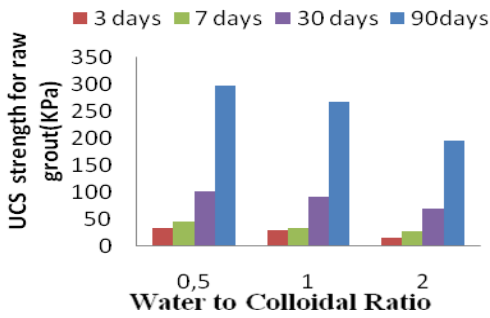


Fig: 7 UCS v/s W/CS Ratio with CaO-2.5%

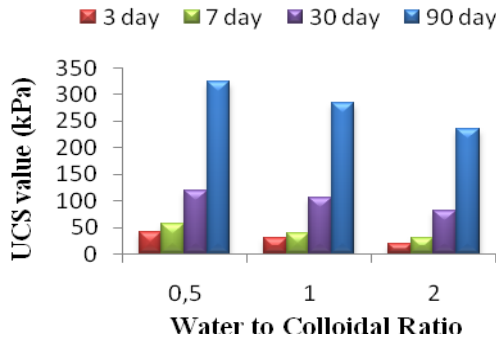


Fig: 8 UCS v/s W/CS Ratio with CaO-2.5% For Raw Grout (Wet Condition) For Grouted sand (Wet Condition)

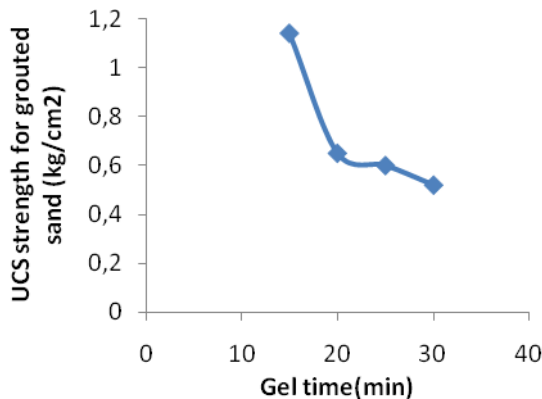


Fig: 9 Gel time Vs UCS strength for sand

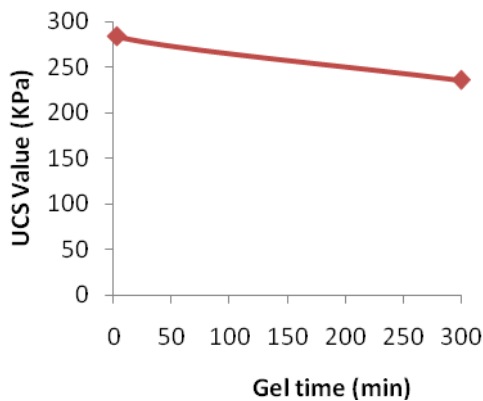


Fig: 10 Gel time v/s UCS strength for grouted sand grouted sand

The relation of gel time and concentration of the Sodium silicate and colloidal Silica grout reactants is studied. It can be conclude that as the concentration of reactant increases the gel time of both grouts reduces, as shown in fig.3 and fig 6. The raw gel strength is increases as the concentration of the reactant increases as the curing time increases as shown in fig.4 and fig.7. The grouted sand strength increases as the concentration increases with increase in wet curing time period as shown in fig. 5 and fig.8. It can be concluded from the study that as relation can be shown from the study that as the gel time reduces with increase in reactant concentration grouted strength increases as shown in fig.9 and fig.10.

VI. SOPE OF THE FUTURE WORK

Trial and error can be performed to find out the optimum grout, and study can be done with laboratory data and practical data. The grout can be design by finding out the different properties and also considering the cost factor, that optimum grout dose can be apply for the solution of practical problem on the field.

REFERENCES

- [1] Caron, C. (1965), "Physico-chemical study of silica gels", Ann. ITBTP-Essais Measures (March), pp.447-485.
- [2] Caron, C. (1967), "Synthetic resins, laboratory tests and application of grouting", Rilem Collague, pp. 183-186.
- [3] Caron, C. (1970), "Grout rheology", Houille Blanche, 192(5).
- [4] Koenzen J. P. (1977), "Time-dependent stress-strain behavior of silicate-grouted sand", Jr. pf Geotechnical Engg. Div. ASCE, Vol. 103, Aug.77, pp. 903-908.
- [5] Shah, D. L. and Shroff A. V., "Field aspects of rheological characteristics of cement and silicate grouts", 51<sup>st</sup> annual technical session R & D session, CBIP.
- [6] Shah, D. L. and Shroff, A. V. (1991), "Rheological characteristics of silicate grouted sand", Proc. 9<sup>th</sup> Asian Reg. Conf. Bangkok.
- [7] Shroff, A.V. et al (1992), "Time-viscosity relationships of Newtonian and Binghamian grouts", Proc. Conf. Grouting, Soil Improvement and Geosynthetics, ASCE, USA, 1:663-676.
- [8] Vipulanandan, C. et al (1992), "Bonding strength of grouts and behavior of silicate grouted sand", Proc. Conf. Grouting, Soil Improvement of Geosynthetics ASCE, USA, 1:700-712.
- [9] Vipulanandan and Krizek J.R. , "Mechanical behaviour of chemical grouted sand", ASCE vol-112; pp-369.



# Mossgrass Seawater Concrete: Comparative Study with tap and Sea Water Concrete

A. Kulshreshtha Yask

**Abstract**—Tap water is always used for construction. Sea water in spite being widely available cannot be used in construction. Use of sea water will cause corrosion in RCC concrete hence durability of structure is reduced. In this paper Mossgrass is used in concrete. Mossgrass is easily available at short distance from sea. Mossgrass is added to sea water and removed after 24 hrs. This water (Mossgrass seawater) is used for casting concrete. For measuring corrosion a new methodology was used which can be useful on site and very economic. Tap water, sea water and mossgrass sea water mixed concrete are compared on basis of compressive stress and corrosion. Results of testing shows the benefits of using Mossgrass seawater mixed concrete over seawater mixed concrete.

**Index Terms**—Corrosion, Mossgrass seawater (M.G seawater)

## I. INTRODUCTION.

Seawater is water from a sea or ocean. On average, seawater in the world's oceans has a salinity of about 3.5% (35 g/L). The average density of seawater at the ocean surface is 1.025 g/ml. Seawater is denser than both fresh water and pure water (density 1.0 g/ml at 4 °C) because the dissolved salts add mass without contributing significantly to the volume.

Seawater should not be used in concrete mixtures because it could contribute to the corrosion of steel reinforcements and inserts. Reinforced concrete structures have the potential to be very durable and capable of withstanding a variety of adverse environmental conditions. However, failures in the structures do still occur as a result of premature reinforcement corrosion. Corrosion of reinforcement has been established as the predominant factor causing widespread premature deterioration of concrete construction worldwide, especially of the structures located in the coastal marine environment. [1]

Corrosion of reinforcement has been established as the predominant factor causing widespread premature deterioration of concrete construction worldwide, especially of the structures located in the coastal marine environment some recent studies seem to suggest that seawater is appropriate mixing water, even though

seawater tends to cause both persistent dampness of the surface of concrete and efflorescence. The presence of chlorides in seawater, particularly sodium chloride, increases the risk of alkali-aggregate reaction. Although the use of seawater can lead to extensive corrosion of reinforced concrete structures, it is misleading to ignore other factors that contribute to concrete deterioration, such as microclimate, air current eddies, solar radiation, aeration, and frequency and duration of wetting. If fresh water is unavailable for curing at an early age, membrane curing should be used as an alternative. [2]

Durability of concrete in a marine environment is a function of its mixture constituents, freeze-thaw susceptibility, abrasion resistance, fatigue strength, and corrosion of embedded metal. [3]

Case histories of deteriorated Portland-cement concretes exposed to sea water, both in mild and cold climates, show that permeability is the most important characteristic determining the durability of concrete. Whether due to improper mix proportions, or poor concreting practice, or cracking of concrete, permeable concretes tend to deteriorate in marine environment. [4]

Due to excessive microcracking, concrete becomes high enough to permit oxygen access to large areas of the reinforcement. Hence electrochemical process occurs resulting in corrosion.[5]

It is confirmed again that between seawater and the constituents of hydrated cement paste, harmful chemical reactions such as carbonation, sulfate attack, and magnesium ion attack can be limited to the surface when well known measures to assure low permeability of concrete have been put into practice. [6]

Fiber reinforced concrete has been shown to have improved fatigue characteristics and improved cracking behavior over conventional concrete. These properties can be advantageously applied to concrete in a marine environment, providing the durability in that environment is satisfactory. [3]

## II. FACTORS AFFECTING CORROSION

Corrosion is defined as ‘the destructive attack of a metal caused by either a chemical or an electrochemical reaction with the various elements in the environment’.

Higher levels of corrosion that occur in the presence of moisture and the hostile gas species, such as chlorine,

ammonia, sulphur dioxide, hydrogen sulphide and oxides of nitrogen, that are often present in the atmosphere.

Corrosion generally depends on nature on material, surface condition, moisture absorption, temperature, humidity and corrosive elements.

But in special case of concrete, the factors that might affect corrosion were evaluated by Song et.al.[7] These factors are:

#### *Environmental effects*

Corrosion itself is the consequence of environmental effects. The seriousness of corrosion of ocean R.C. structures is therefore directly related to the location of the structural members. Atmospheric corrosion, however, is less serious and underwater corrosion is the lightest.

#### *Construction quality*

It is very important to ensure construction quality for preventing the R.C. structures from corroding. Some general problems of construction quality include: the thickness of concrete cover is not enough; the joints between beams and slabs are not constructed carefully, or the concrete is not compacted properly and the reinforcement is not placed correctly.

#### *Thickness of the concrete cover*

The thickness of the concrete cover plays an important role in preventing the corrosive substance from reaching the surface of the reinforcement. The results of the test show that when cover thickness is increased from 3 to 4 cm, both the rate of weight loss and the rate of rusting could be reduced about 91% after six cycles of wetting and drying.

#### *Property of the concrete material*

The best protection against corrosion of reinforcement is to use well-compacted concrete, while the density of concrete is mainly affected by several factors: the water cement ratio, the content of the cement, the size of the aggregate and the quality of the water reducing agent. Increasing the content of cement can improve the resistance to corrosion. Regarding the effect of aggregate size, the larger the size of the coarse aggregate, the more serious the corrosion of the reinforcement. This is because the non-uniform shrinkage cracks are formed easily in concrete with larger size of the coarse aggregate.

The most important causes of corrosion initiation of reinforcing steel are the ingress of chloride ions and carbon dioxide to the steel surface. After initiation of the corrosion process, the corrosion products (iron oxides and hydroxides) are usually deposited in the restricted space in the concrete around the steel. Their formation within this restricted space sets up expansive stresses, which crack and spall the concrete cover. This in turn results in progressive deterioration of the concrete. As a result, the repair costs nowadays constitute a major part of the current spending on infrastructure. Quality control,

maintenance and planning for the restoration of these structures need non-destructive inspections and monitoring techniques that detect the corrosion at an early stage. [8]

### III. MOSS GRASS

Mosses are small, soft plants that are typically 1–10 cm (0.4–4 in) tall, though some species are much larger. They commonly grow close together in clumps or mats in damp or shady locations.

Since mosses have no vascular system to carry water through the plant, they must have a damp environment in which to live, and a surrounding of liquid water to reproduce. And since mosses are photosynthetic, they require enough sun to conduct photosynthesis. Moss species can be classed as growing on: rocks, exposed mineral soil, disturbed soils, acid soil, calcareous soil, cliff seeps and waterfall spray areas, streamside's, shaded humus soil, downed logs, burnt stumps, tree trunk bases, upper tree trunks, and tree branches. Moss is available in large quantities. It is also cheap to extract it from its habitat.

Moss grass had been used by tribal people for bedding, basketry and wound dressing. Food storage baskets and boiling baskets were also packed with mosses. The uses for intact moss are principally in the florist trade and for home decoration.

Decaying moss in the genus *Sphagnum* is also the major component of peat, which is "mined" for use as a fuel, as a horticultural soil additive, and in smoking malt in the production of Scotch whisky. *Sphagnum* moss is harvested while still growing and is dried out to be used in nurseries and horticulture as a plant growing medium. The moss retained large volumes of water which helped extinguish the flames.

### IV. OBJECTIVE

It is a well known fact that we should use that water for construction which we can drink. But these days world is facing water scarcity in terms of drinking water. A large amount of water is used in construction hence drinking water availability is lower than its source. In future a time may come where we don't have enough water for construction and drinking. So the only available option is Sea water. It is not used in concrete mix because it leads to rapid and extreme corrosion.

In this paper an attempt for making a concrete which uses sea water and don't corrode that fast and extreme. Moss grass is used in making such a concrete mix.

#### *Why Moss Grass?*

Moss grass is easily available near sea shore. The construction which uses sea water in concrete mix will essentially be located near sea shore. Onshore structures

are constructed by using sea water made drinkable using reverse osmosis.

Moss is having lipids. It can absorb a large amount of water. It can be assumed that when it absorbs water, the minerals and salt in water will be left behind on its surface or skin. Later when moss grass is removed from salty water, some chlorides which are on its skin are removed. If salt content is low in water then corrosion will be less and not be very fast.

V. METHODOLOGY

In the experiment we have to make samples of concrete, reinforced concrete and test it for strength and corrosion. For compressive strength digital compressive testing machine is used. For testing corrosion different methods are available. But these methods are costly and need sophisticated instruments. Instruments which are portable are also not easily available. In that case there is a need of some method which uses the easily available material and can be assembled quickly. Such a method is discussed here.

Corrosion test

This test is performed on concrete cylinder (150mm diameter and 150mm height) having a 12mm reinforced rod. Materials required are non-insulated copper wire, high precision digital multimeter and insulating tape or tight rubber band.

Copper wire is cut into 12 small pieces of 30mm each. These copper wires with help of insulating tape or rubber band is tightly made in contact to concrete surface. The whole configuration is as shown:

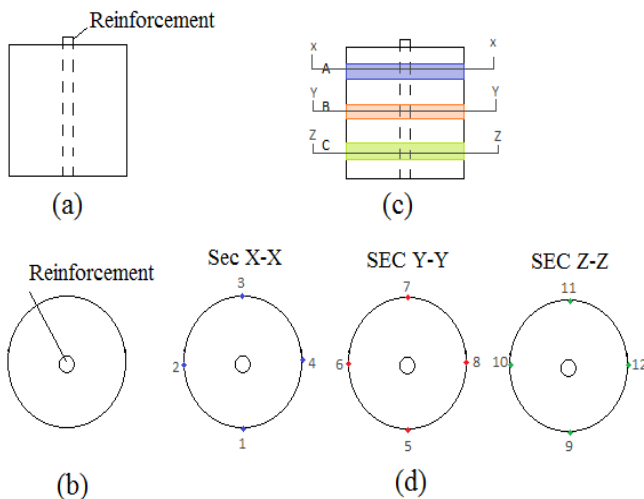


Fig.1. (a) Section of concrete with reinforcement ; (b) Top view of concrete with reinforcement at centre; (c) Concrete with insulated tape over half piece of copper wire; (d) view at different sections, dot show the position of copper wire piece.

The copper wire is kept half exposing the atmosphere and half under the insulator. Multimeter is having two terminals. One terminal is connected to one of these

copper wires and other on the reinforcement as shown in Fig.3.2. For better contact copper and zinc wire is connected to reinforcement rod.

Thus voltage reading is shown in multimeter. Similarly potential difference between reinforcement rod and each of 12 points (having one copper wire piece each) is measured and contour graph can be plotted. More the voltage is indicated more is the corrosion.

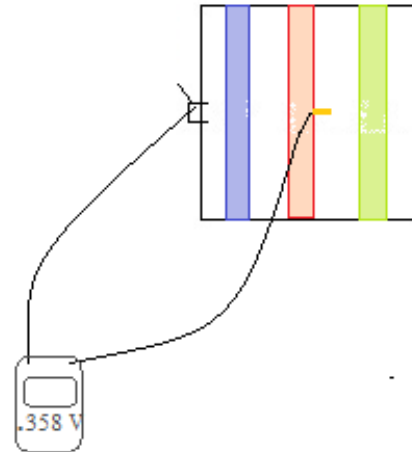


Fig.2 terminals of multimeter connected to reinforcement rod and copper on surface of concrete respectively

VI. OBJECTIVE

Preparation of moss grass sea water

Moss grass was collected and about 150gm was kept in a bucket containing 6 litres of sea water. The moss grass was taken out from the bucket after 2 days (Fig.4.1). Water in bucket is then passed through fine sieve so that no floating grasses exist in it. This moss grass sea water is then used directly in casting of concrete.



Fig.3 Moss grass taken out of sea water

Casting and testing of cubes

3 different type of water was available for casting; tap water, sea water and moss grass sea water. 6 cubes of size 150mm × 150mm × 150mm were casted for each kind of

water sample. So a total of 18 cubes were casted. Concrete mix design of M25 was used for experiment. PPC cement was used in concrete. All the samples were cured in saline water.

These cubes were tested for compressive strength on 9<sup>th</sup> and 28<sup>th</sup> day.

*Casting of cylinder with central reinforcement*

3 cylinders of 150mm diameter and height 150mm for each kind of water sample were casted with central reinforcement of size 12mm. These cylinders were casted in special mould capable of holding reinforcement rod. A total of 9 such cylinders were casted and cured in saline environment.

*Testing of corrosion*

Corrosion testing was done by an own adopted methodology as explained earlier. Fig.4.2 shows the actual setup used in corrosion testing. A,B and C are three strips of insulator on which 4 copper wire piece at equal distance is attached.



Fig.4. Corrosion testing setup

Terminal of multimeter is connected to the copper wire attached on concrete surface via insulating tape as shown in fig.4.3. Rubber band can also be used instead of insulating tape.



(a)



(b)

Fig.5 (a) copper wire attached on concrete surface through insulating tape; (b) terminal/clip of multimeter connected to copper wire.

VII. RESULT AND DISCUSSION

Compressive stress analysis

CONCRETE	COMPRESSIVE STRESS - 7 DAYS	COMPRESSIVE STRESS - 28 DAYS
TAP	27.4889	39.4415
SEA WATER	33.7970	44.7066
M.G SEA WATER	32.4770	40.9437

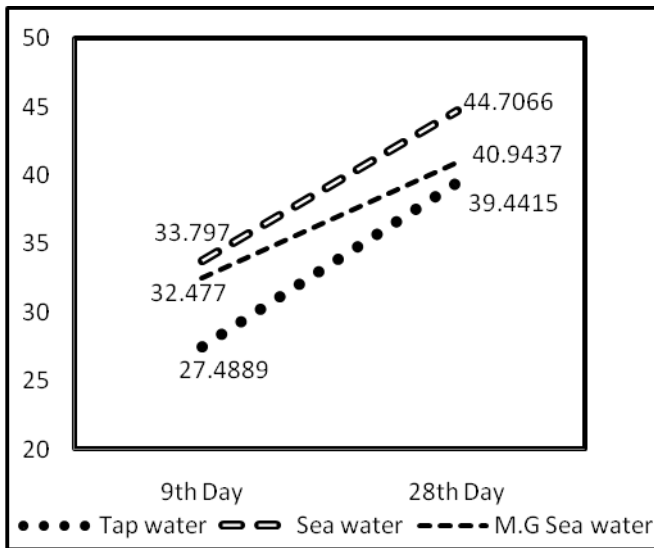


Fig 6. Comparison of compressive strength

It can be seen that sea water gain more strength initially. Strength of tap water is low compared with sea water. Moss grass sea water is having strength in between other two. Strength of M.G Sea water seems to be moving close to tap water concrete strength with time. It can be stabilized that strength of all concrete will be approximately same after some years. The reason for fast gain of strength in case of sea water is due to faster alkali-aggregate reaction. Sea water is having 15% and 10% more strength than tap water and moss grass sea water respectively.

Comparative studies of different samples at each section

POINTS	TAP WATER (VOLT)	SEA WATER (VOLT)	M.G SEA WATER (VOLT)
1	0.41	0.517	0.461
2	0.425	0.551	0.438
3	0.419	0.513	0.47
4	0.417	0.518	0.443

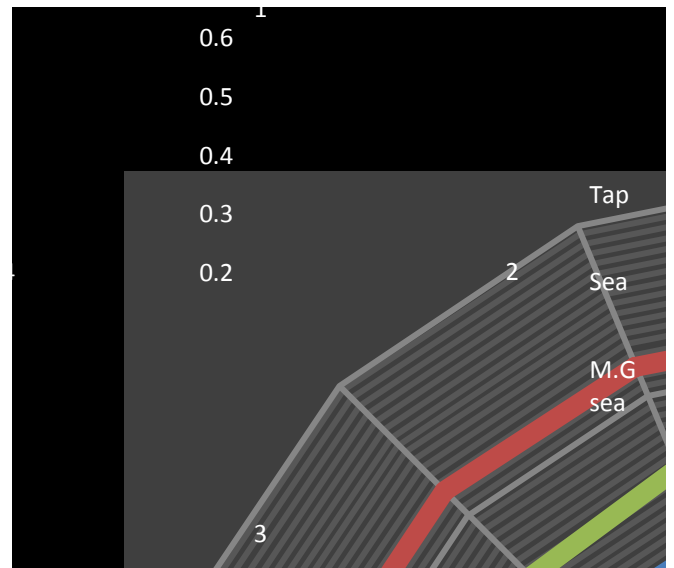


Fig 7. Comparison at section-A

The corrosion result on 28<sup>th</sup> day testing is shown here. Corrosion on 3<sup>rd</sup> day were also tested. Corrosion is increasing with time. Section A is much closer to the top of concrete cylinder where a small part of reinforcement is extending from concrete which is exposed to atmosphere. Hence this section will have more corrosion and chlorine from saline water can penetrate easily to this section. It can be seen that voltage of seawater > Mossgrass sea water > tap water. Hence corrosion of sea water is highest, whereas corrosion of Mossgrass seawater is moderate. This indicate that Moss Grass seawater is more suitable for concreting.

POINTS	TAP WATER (VOLT)	SEA WATER (VOLT)	M.G SEA WATER (VOLT)
5	0.361	0.453	0.385
6	0.346	0.465	0.385
7	0.365	0.47	0.382
8	0.385	0.437	0.389

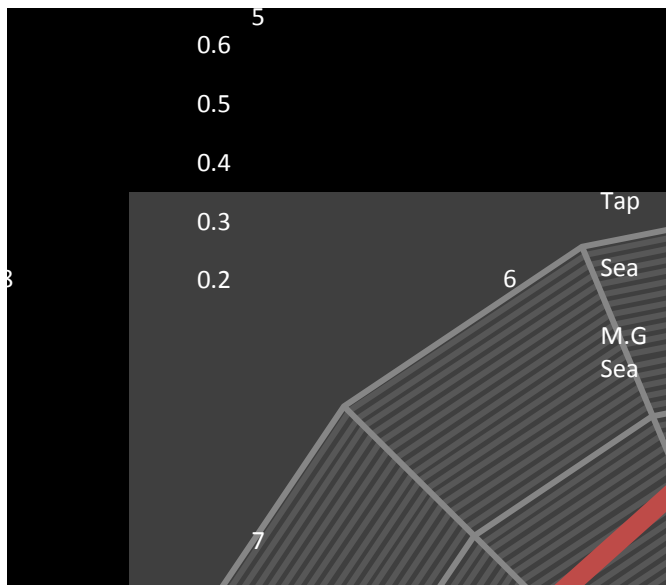


Fig 8. Comparison at section-B

Section B is at the middle and it will take a long time for chlorine to penetrate it. Hence corrosion at this section is less. Here potential difference in case of tap water and moss grass sea water is almost same, which ensure better effect of M.G sea water over sea water.

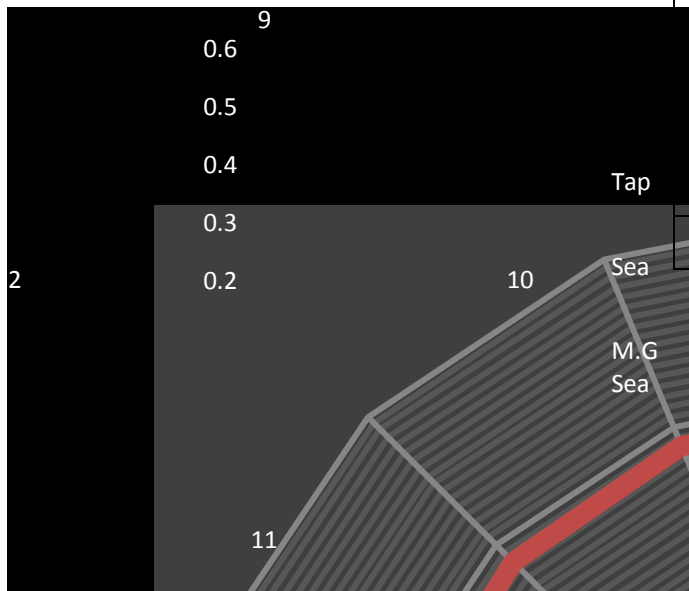


Fig 9. Comparison at section-C

Section C is near bottom of concrete, so similar conditions to section A can be seen. Sea water mixed concrete has more corrosion then Moss grass seawater concrete.

It can be analyzed that in all cases corrosion in sea water is highest. Corrosion in Moss grass sea water is moderate and corrosion in tap water is low.

We can say that; corrosion rate: sea> moss grass  
sea> tap

ACKNOWLEDGEMENT

I would like to thank Prof R.S Jangid, department of civil engineering, IIT B, who guided me throughout my project. I would also thank Snehal mevada sir, a PhD student under professor R.S jangid who always provided me with necessary knowledge and material to carry out my experiment. I also thank Pradeep desai, project staff at CESE, IIT Bombay for providing me sea water samples. My sincere thanks to Structural evaluation and material testing (SEMT) technologies laboratory for their support in preparation and testing of Cavity wall. I thank KVPY for providing me funds for successful completion of project.

VIII. CONCLUSION

- 1) Sea water mixed concrete shows gain in strength in earlier days compare to tap water mixed concrete. Sea water mixed concrete has more strength than tap water concrete but after long duration (in

POINTS	TAP WATER (VOLT)	SEA WATER (VOLT)	M.G SEA WATER (VOLT)
9	0.392	0.491	0.398
10	0.371	0.48	0.422
11	0.377	0.484	0.49
12	0.386	0.482	0.417

decades) both will have approximately same strength.

- 2) Moss grass sea water mixed concrete also show gain in strength in earlier days. But its strength after some days (approx. 1 month) is comparable to tap water concrete.
- 3) Corrosion measured after a month shows that it is higher in sea water mixed concrete, moderate in moss grass sea water mixed concrete and very low in tap water mixed concrete when compared.
- 4) Corrosion depends on chloride ions and carbon dioxide, temperature, humidity, surface condition, moisture absorption, construction quality and property of the concrete material.
- 5) Rate of corrosion is also high in sea water mixed concrete, moderate in moss grass sea water mixed concrete and low in tap water mixed concrete when compared.
- 6) Moss grass sea water mixed concrete is a concrete whose properties lies between tap water and sea water mixed concrete.

- 7) Therefore it can be used in places where low reinforcement is used. It will be more durable as compared to sea water concrete.

#### REFERENCES

- [1]Raupach, M., Schiebl, P. "Macrocell sensor system for monitoring of the corrosion risk of the reinforcement in concrete structures" NDT&E International, 34(6), pp.435-442,2001.
- [2]Neville,A. "Seawater in the mixture", *Concrete international* 23(1), pp.48– 51,2001
- [3]Hoff,G.. "Durability of fiber Reinforced Concrete in a Severe Marine Environment", Special Publication-ACI 100, pp.997-1042,1987.
- [4]Mehta, P. "Durability of Concrete in Marine Environment-A Review", Special Publication-ACI 65,pp.1-20,1980.
- [5]Mehta, P. and Gerwick Jr.,"Cracking-Corrosion Interaction in Concrete Exposed to Marine Environment", *Concrete International* 4(10), pp.45-51,1982.
- [6]Mehta, P. "Durability of Concrete Exposed to Marine Environment-A Fresh Look", Special Publication-ACI 109, pp.1-30,1988.
- [7]Song,Y., Song,L. and Zhao,G. "Factors affecting corrosion and approachesfor improving durability of ocean reinforced concrete structures", *Ocean Engineering*, pp.779–789, 2004.
- [8]Song, H. and Saraswathy, V. "Corrosion Monitoring of Reinforced Concrete Structures – A Review", *International Journal of Electrochem. Sci.*, Vol. 2, pp.1-28,2007.

# Cooling through cavity walls

A. Kulshreshtha Yask

**Abstract-** Cavity wall construction is an old construction technique, widely used in cold areas. Cavity wall construction has many advantages like – Damp proofing, thermal insulation, fire resistance, sound proofing and low load on foundation. Cavity is filled by different material for thermal insulation. Soil as filler is used in cavity walls. Soil is cooled by constant dripping of water through earthen pot. Waste water from terrace garden can also be utilized for this purpose. Effect of the cooling is studied on walls. No electricity is consumed in this process of cooling. This technology is energy efficient and make house sustainable.

**Index term-** Cavity wall, Cavity wall insulation,

## I. INTRODUCTION

CAVITY walls consist of two vertical masonry separated by an air space and joined together by metal ties. Cavity wall construction is widely used in United Kingdom and U.S.A, but in India it is not practiced well. The cavity is generally 4cm to 10cm and the thickness of walls is ranging from 30cm to 45cm. Generally cavity wall is having air in the cavity. As to increase the thermal insulation efficiency of cavity wall now a day's many filler material is added to it like: White and yellow wool, polyester balls [1]-[4]. Soil is available everywhere on earth. In this model soil is used as filler in cavity wall. It is available easily and satisfies the property of filler we need for design. We can construct brick masonry on either side of cavity filled with soil or we can construct brick masonry and concrete on either side. We generally prefer brick on outer side and concrete side facing indoor. Hollow concrete blocks are also used in cavity wall construction. Hollow concrete blocks are mostly used as they have better thermal insulation and total load on foundation is reduced.

There is a need to provide a damp proofing coating on the cavity side of inner wall. If moisture from soil is penetrated to wall then it can cause strength loss, efflorescence etc.

This will ultimately damage wall. Initially cost of construction will rise due to the given requirement but its future saving is more. In this method, we have to pour water throughout the cavity.

We should always remember water should be as cool as possible so as to get best results. We don't need large amount

of water as we have to do dripping. Water can be taken from overhead tank or we can utilize the waste water in this technique. Water is removed out from bottom through pit holes, which are drilled in concrete. We can use earthen pots drilled with a small hole at bottom. Water through pipe can be poured in all earthen pots. Water can be cooled to 3-10 degree less than normal atmospheric temperature.

Due to limitation of space concept of terrace gardening is getting global acceptance. Many people grow plants in pot and place them in terrace. Watering is important for plant. Pots have holes at their bottom through them through which water excess water comes out. This water can made to enter cavity walls through small holes drilled on terrace. When water comes in contact with soil it releases heat. It is important to release this heat, which can be removed through these holes. So these holes serve two purposes. Cavity wall reduces heat conduction up to 60%. By using this method it can be reduced to more high value.

## II. EXPERIMENT

An experiment on cooling through cavity walls was conducted. The location for experiment was Structural evaluation and material testing (SEMT) technologies laboratory, IIT Bombay. A brick wall of height 40cm and width of 68cm was constructed. The size of brick used is 22\*11\*7. By leaving a space of 10 cm a wall of precast light weight concrete was placed (Fig1.) Bricks were joined by mud mortar (Fig 2). Average quality of bricks was used in the experiment. A thermocoel sheet was also placed in cavity in contact with brick masonry.

The soil was filled in different layers as shown in Fig 3. This type of pattern is also seen in traditional water purification through earthen pot. Jute bag strip is also placed above fine soil layer. Jute bag is waste material for many shopkeepers and easily available. Jute is good absorber of water. High moisture is accumulated in soil near to jute bag strip due to which soil in this area will be cooler. We can arrange different jute strip keeping a constant distance between them. Wall ties should be painted with water proof paint. Jute bag should not be kept near to wall ties else the chances of corrosion increases.





Fig.1 Top view of a section of cavity wall constructed



Fig 2. Side view of a section of cavity wall constructed

FINE SOIL+ COARSE SAND + FINE SAND	9 cm
FINE SOIL	8 cm
COARSE SOIL	6.8 cm
JUTE BAG STRIP	0.2 cm
FINE SOIL	8 cm
COARSE SOIL	8 cm

Fig 3. Layer of soil placed in cavity of cavity wall

Earthen pot was used in experiment as it was easily available within affordable price limits. It can cool water in some time. Desert water bag was also used which can cool water in a short interval of time. However it is difficult to find out exact time T in which water cools down to 1 C, approximately in 5 min to 10 min water temperature falls 1 C. Water is constantly dripping out of earthen pot when it is filled it with water. Dripping rate is very slow. Position of earthen pot was changed timely so as to moist soil uniformly. Even water can be poured in large quantity at one moment of time, so there will be no need of adding water continuously. The surface on which wall is constructed is made up of concrete having a gentle slop. Excess water comes out through one side.

III. RESULTS

For measuring temperature we have to use an accurate and precise instrument so I used digital thermocouple having Cr-Al. It can measure up to a high temperature of 1200 C.

Temperature readings (T) were taken for different time and date. Temperature of atmosphere, water inside earthen pot, water inside desert water bag, water inside the pond which is source of water which is used for pouring in soil , moist soil in cavity, aluminium foil in cavity wall, brick masonry exposed to sun ,concrete facing indoor and concrete exposed to sun is measured and result is tabulated below as shown in Table 1, Figure 4, Table 2 and Figure 5 :

TABLE 1. TEMPERATURE MEASURED ON A SUNNY DAY

	Temperature (°C)				
	12:30 hr	14:20 hr	15:00 hr	16:00 hr	18:45 hr
Atmosphere	34	37	35	33	29
Brick	34	36	35	34	30
Concrete	32	35	34	33	30
Exposed concrete	40	44	42	39	32
Water in Earthen pot	26	28	27	26	26
Water in water bag	26	27	28	27	26
Water in pond	31	33	31	30	27
Soil	32	29	29	28	27
Aluminium foil	31	29	29	29	28

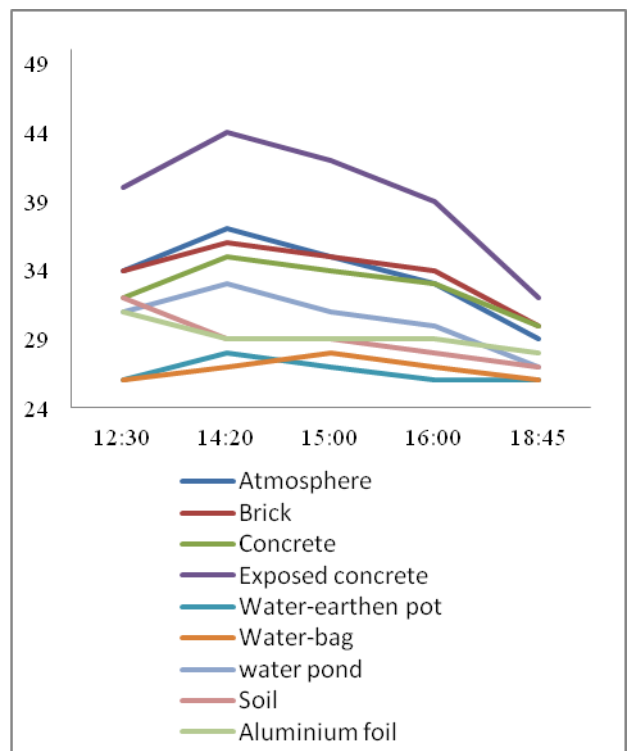


Fig.4. Graph of temperature and time on a sunny day

On a sunny day, by referring the graph we can interpret useful information:

1. The atmospheric temperature rises in morning and reaches its peak afternoon in around 2 pm. This is the time we need cool or normal environment for working efficiently. Temperature decreases after the peak time is almost uniform.
2. Brick is exposed to sunlight. It has low thermal conductivity. Temperature of brick under shadow is low, so we can plant trees near our house for receiving shadow.
3. With increase in atmospheric temperature, temperature of concrete wall on a side of cavity increases. As this wall is not in direct contact of sun so its temperature is less than atmospheric temperature. Concrete is having a conductance of 1.7 W/(m.K)
4. The surface of concrete in contact with direct sunlight becomes very hot and it can reach up to 50 C temperatures. While concrete conduct low heat through its thickness.
5. Temperature of water inside earthen pot and desert bag is 5 to 10 degree C lower than atmospheric temperature. When the wind is flowing with high velocity then cooling is faster. Temperature drop of 5 C is achieved in 30 min provided other conditions are normal. The water in pot and bag is water taken from pond having temperature of water 2-3 degree C less than normal atmospheric temperature.
6. Powai soil was used in this experiment. When water is poured in soil, the heat of soil is released. Its temperature decreases significantly .temperature of soil depend on temperature of water poured and type of soil. Temperature of soil is generally near 28 to 29 degree C. So the temperature of cavity is much lesser then atmospheric temperature.

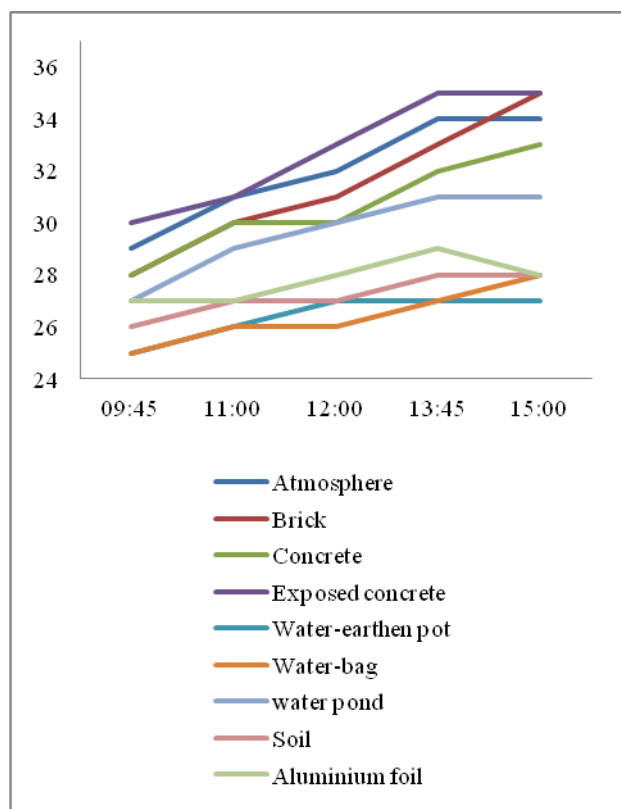


Fig 5. Graph of temperature and time on a sunny and cloudy day

On a sunny and cloudy day, by referring the graph we can interpret useful information:

1. The atmospheric temperature rises in morning and it is almost constant till evening due to clouds coming in sky. Temperature decreases very steadily. Temperature variation for whole day is small.
2. Brick is not exposed to sunlight for greater time. It has low thermal conductivity. Temperature of brick doesn't increase too much.
3. With increase in atmospheric temperature, temperature of concrete wall on a side of cavity increases. As this wall is not in direct contact of sun so its temperature is less than atmospheric temperature. Concrete is having a conductance of 1.7 W/(m.K). Concrete conduct low heat through its thickness.
4. Temperature of water inside earthen pot and desert bag is 3-7 degree C lower than atmospheric temperature. During cloudy day generally wind flow with high velocity. When the wind is flowing with high velocity then cooling is faster. The water in pot and bag is water taken from pond having temperature of water 1-3 degree C less than normal atmospheric temperature.
5. Temperature of soil depends on temperature of water poured and type of soil. Temperature of soil is generally

TABLE 2.

TEMPERATURE MEASURED ON A SUNNY AND CLOUDY DAY

	Temperature (°C)				
	9:45 hr	11:00 hr	12 hr	13:45 hr	15:00 hr
Atmosphere	29	31	32	34	34
Brick	28	30	31	33	35
Concrete	28	30	30	32	33
Exposed concrete	30	31	33	35	35
Water in Earthen pot	25	26	27	27	27
Water in water bag	25	26	26	27	28
Water in pond	27	29	30	31	31
Soil	26	27	27	28	28
Aluminium foil	27	27	28	29	28

near 27 to 28 degree C. So the temperature of cavity is much lesser than atmospheric temperature.

#### IV. ELECTRICITY GENERATION FROM CAVITY WALL

It was very interesting to discover that an electromotive force (current) and voltage was generated in cavity. It was found that if adequate moisture is maintained in soil current and voltage can be made nearly constant. This can be done with help of constant dripping. We were already using dripping in soil. So we can use same soil for generation of electricity.

Copper tube and aluminium foil was used as electrode. Various experiment were done to increase the current and voltage output. I was able to generate voltage of 1.2V and corresponding current of 800 microampere. By using marine clay current of 7 mA and corresponding voltage of 0.7 V was generated. Clay is rich in salt content. Salt increases the conductivity. Hence there is a possibility of generation of electricity if this type of setup is arranged in series. This electricity can be used in lighting LEDs during power cuts.

#### V. OTHER METHODS OF COOLING THROUGH CAVITY WALLS TESTED

Other methods were also tested which can be useful for cooling houses. These setups will generally be made in between lintel and roof level.

##### A. Cooling through cavity walls using porous pipe, jute bag and soil

Cool air is denser than warm air so it will ultimately settle down so there is no need to fill cavity with soil everywhere for better effect. Porous pipe is a special type of pipe basically designed for irrigation purpose. This pipe is having many micro pores through which water comes out in form of droplets. Speciality of this pipe is that we can use it easily below soil and it produces uniform dripping all around its surface. Porous pipe receive cool water from earthen pot and it releases maximum of its water in to the soil (on one side) and jute bag on other side. We have keep soil because it will always remain damp and jute is having a property to absorb moisture, which make jute filled with moisture. These cool water droplets are absorbed by the air circulating due to the fan and hence cool the room. A gentle slope towards the cavity is provided on the lower side of opening so that droplet always falls in cavity. Water is guided out through pit holes drilled in concrete. We can even fill the lower portion of cavity with soil. This design is economical, relatively easy to construct and over all loads on foundation will also reduce.

##### B. Cooling through cavity walls using copper hollow tube

By blowing air from one side to other side, a decrease of 1-5 degree C in temperature is observed near copper tube end enclosed in a chamber. In this design circulation inside the room take place naturally due to velocity of wind. This system will surely help those floors where wind flows with sufficient velocity. Air will enter in tube, get cool in cavity and enter inside room. The length of path and velocity of air in it determines the cooling caused inside the room. We can use a special flap on the opening of tube facing atmosphere which restrict large quantity of air to come in at specific time.

##### C. Cooling through cavity walls using pipe dripping in an opening

It is a modification of traditional cooling in which village people firstly makes jute bag wet and then hang it over door or window so that the air coming in become cool. But they have to wet jute bag after some time when it losses considerable amount of moisture.

In this design an opening like a ventilator is constructed. This opening is made both in brick masonry and concrete. This construction should be done above lintel level. This design is useful for upper floors where wind flow at high velocity.

(a) Root bag - Temperature was drop to 3 degree C when wind passed through the wet bag. The amount of air passing through is relatively lesser than grass and root bag.

(b) Dry grass bag - Temperature was drop to 2 degree C when wind passed through the wet bag. The amount of air passing through is relatively higher than root bag and lower than jute bag.

(c) Jute bag - Temperature was drop to 2 degree C when wind passed through the wet bag. The amount of air passing through is relatively higher than grass and root bag.

#### VI. CONCLUSION

A part of wall is constructed so the result of cumulative walls enclosing an environment will be different. It will surely be more effective. This idea is new and can be implemented when workmanship is excellent. Need of man hole is also needed in case of any failure. But surely this is eco-friendly and energy saver. We can conclude:

1. This system is not releasing any type of harmful waste, so it is eco friendly. Even waste material can be used in this system.
2. Whole world is focusing on energy conservation. The amount of electricity consumed by A.C and cooler for cooling a building is very high. This way is totally mechanical and not consuming any electricity.
3. It may cost high at initial stage, but it is economical for long run.

4. Due to limitation of space, roof gardening is becoming popular. The waste water of gardening can be utilised in this system.
5. Electricity can also be generated when soil battery is connected in series. This electricity can be used during power cut.
6. It makes a house or a building self sustainable.

#### REFERENCES

- [1] Punmia,B.C. et.al.,2008. *A text book of Building construction*. New delhi: laxmi publication ltd.
- [2] Masonry advisory council, Illinois.,2002. *Design Guide for Taller Cavity Walls* [online]. Available from: [www.maconline.org](http://www.maconline.org). [Accessed 19 June 2010]
- [3] TESCO.,2010. *Cavity wall insulation for wall to wall energy saving* [online]. Available from: [www.tescohomeefficiency.com](http://www.tescohomeefficiency.com). [Accessed 19 June 2010]

# Behavioral study of commuter for mass transit system – A case study of Ahmedabad BRTS

A. Vishal P. Patel

**Abstract:** Urban transportation system in metro cities become a challenge to the authority by various means named as traffic congestion, delay, parking, accidents, noise and air pollutions. In response to the challenge, various public transport systems are initiated by authorities. Few of them are local bus services, Bus Rapid Transit System (BRTS), Metro Rail, Tram and light rail etc. Ahmedabad Municipal Corporation (AMC) adopted BRTS to strengthen the present public transportation system of Ahmedabad Municipal Transport Service (AMTS) since 2009. Behavioral aspects of commuter allied to income, trip purpose, age, sex, previous mode of transportation etc. are studied in present work and reported. Experience sharing of commuter is also addressed to understand the performance of BRTS.

**Index Terms**—Behavioral, BRTS, Commuter, Public Transport System,

## I. INTRODUCTION

URBAN transportation system plays a vital role in day to day life of citizens. Movement of people is performed for various purposes using various modes of transportation system. Movement of people is accompanied by the movement of goods to satisfy the daily need of human being results into the additional demand of transportation infrastructure. Purpose of trips may vary from education to work, social to recreational or return trips. Modes of transportation ranging from the road transport to rail transport and if possible water ways also. In the initial days of urbanization, development of infrastructure facility took place with the vision of next 30-35 years. At the end of planning period (30-35 years), new generation faces problems due to the gap between demand and capacity of infrastructure. Known statement “**Necessity is the mother of invention**” is also applicable here in terms of advancement of technology and policy to cater the demand over the capacity. Capacity of road is defined in number of vehicles those can pass through specified width of road in a unit time. Vehicle’s functional characteristics and its size affects significantly to the capacity of roads and thus all types of vehicles should be converted to the same platform using the concept of “**P**assenger **C**ar **U**nit”. Introduction of public transportation system aims to reduce the number of vehicles from the road without compromising the number of passenger

movements in a unit time. Other associated advantages can be summarized as improvement of speed and optimized travel time, reduction in pollution, saving in fuel consumption, safe movement, reduction in parking requirement etc. Looking to the importance of public transportation system in urban areas, usage of the same should be promoted for the realization of benefits.

Ahmedabad Municipal Transport Service (AMTS) is one of the oldest public transportation systems in Ahmedabad city since independence in 1947. Area of AMC also increased from 52.5 square kilometer to 475 square kilometer after independence. During same period, population of Ahmedabad city also crossed to 5.5 millions. Considering the span of area and increased population, necessities of mass rapid transit system is felt by authority. In the result, BRTS is introduced in year 2009 to provide a fast, safe, comfortable and economic journey with the dedicated right of way on a road. In present study, an attempt is made to understand the attributes associated with the users of public transportation system. Mainly age, sex, trip purposes, previous modes of transportation, Income, vehicle ownership etc. are considered. Co-relation of such attribute, leads to the probabilistic analysis of parameters for modal shift after development/improvement of public transport system. Compliments/complains for the BRTS are also collected in the study to have feedback for the performance.

## II. DATA BASE

### A. Survey Methodology

Based on the literature review, Revealed Preference Survey is preferred to carry out at BRTS bus stations as well as during journey in BRTS. Secondary data source from the competent authority and published reports are used to validate the data as and when required. Considering the limitation of number of surveyor, questionnaire format is designed in a manner to collect all the possible information in a short duration. To capture all type of possible commuters, choice of BRTS station and time period of survey is kept accordingly. Survey was scheduled for weekdays only to keep data generalized in nature. For reference, filled questionnaire format is given below.

**Questionnaire for modal shift analysis**

QUESTIONNAIRE NO - ..... 8

Gender	Age	Monthly income	Earlier mode of transportation	Trip purpose
M	23	-	01	02

Origin	Destination	Vehicle ownership	Travel time in bus	Approx. distance
Anjali	Nehru nagar	1		

**Codes used for earlier mode of transportation**

AMTS = 01	Auto rickshaw = 04
Two-wheeler = 02	Jeep = 05
Car = 03	

**Codes used for monthly income**

<5000	5000-10000	10000-20000	20000-50000	>50000
1	2	3	4	5

**Codes used for trip purpose**

Work = 01	Education = 02	Others = 03
-----------	----------------	-------------

Number of people in bus : ..... 70

Comments :  
 i) Crowded  
 ii) Buses to be increased  
 iii) Not sure about MRTS

Fig. 1. Questionnaire Format

**B. Data Presentation**

Total about 886 survey respondents are collected to further analysis. Out of these 886, only 7 respondents are shifted from the cycle/walk/jeep to the BRTS which are neglected in all further analysis. So ultimately, data compilation, presentation and analysis are carried out for 879 respondent's data using Microsoft Excel. Range for the age and income is derived on a logical ground and corresponding coding is used for further analysis.

**III. ANALYSIS**

Analysis is carried out to know the % shift from various modes to BRTS, % share of trip purpose, effect of vehicle ownership and monthly income on commuter's behavior and in the last combined analysis based on findings of the each analysis.

**A. % Shift from various Modes to BRTS**

Data shows that 43% commuters are shifted from one public transportation system (AMTS) to other (BRTS). So the benefit of shifting for this category of commuter to the overall transportation system is negligible. At the same time % shift from private modes is around 57%, which positively contribute towards the relief on urban roads from traffic congestion and pollutions.

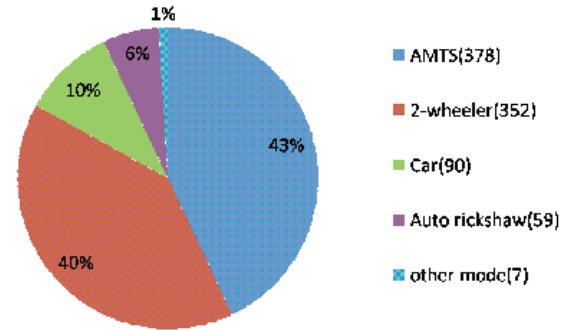


Fig. 2. % Shifts from various Modes to BRTS

**B. % Share of Trip purpose**

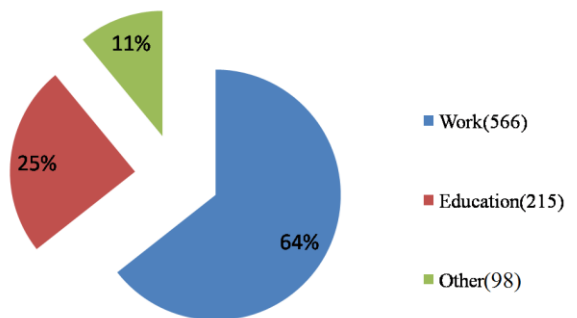


Fig. 3. % Share of Trip Purpose

It seems that BRTS is predominantly utilized for work and education purpose. Work contributes around 64% of total trips and 25% of the total trips performed for education purpose. Other trips count the social and recreational/ entertainment purpose.

**C. Vehicle Ownership**

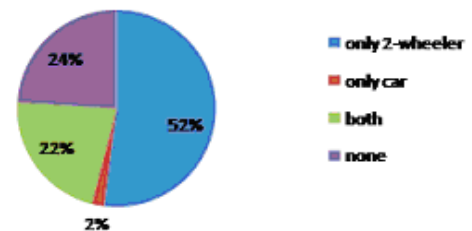


Fig. 4. % Share Trips by Vehicle Ownership

54% BRTS commuters have their own two-wheeler or car. 22% commuters have car and two-wheeler both. This reflects the increase in vehicle ownership with rapid urbanization. Use of BRTS in spite of own vehicles may be contributed due to the inflation.

D. Monthly Income

It is quite surprising that nobody from the surveyed respondents have their monthly income above Rs.50,000.

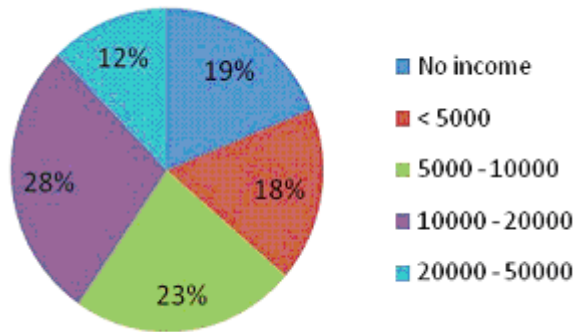


Fig. 5. % Share Trips by Monthly Income

High vehicle ownership compared to monthly income is a prime factor contributing to the traffic congestion on urban roads.

E. Income Vs. Trip purpose

Out of total 65% work trips, 54% contributed from the commuters having monthly income less than Rs. 20,000. It indicates that in spite of safe and faster movement through BRTS corridor, attraction from higher income group commuter is limited. Absence of proper feeder services and overcrowded buses during peak hours are the main reasons for the same.

Education trips with income group hints for the community of part time job holder.

F. Age group Vs. Trip purpose

TABLE 1  
Trip Distribution (Age group Vs. Trip purpose)

Purpose of Trip	Age Group (years)			
	<25	26-40	41-58	>58
Work	19.08	53.18	26.85	0.89
Education	95.81	2.79	0.93	0.47
Other	11.22	24.49	55.11	9.18
	<b>36.97</b>	<b>37.66</b>	<b>23.66</b>	<b>1.71</b>

73% of work trips are contributed from the commuters having age less than 40 years. For Education Purpose, trips in age group above 26 years contributed from the part time learner mainly during evening hours. Significant relationship between age group and trip purpose can be established by adding income as an indicative parameter.

So analysis of three parameters named as trip purpose (work only), age group and income group is given here.

Consider commuters having age up to 25 years, majority (81%) of commuters going for work trips belong to monthly income less than 10,000.

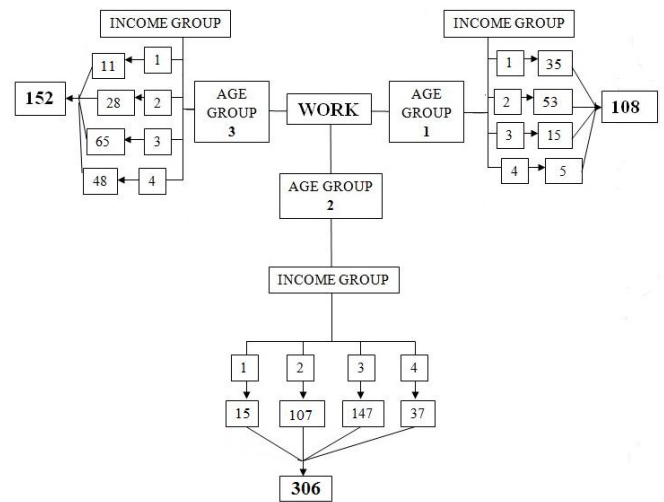


Fig. 6. Effect of Age and Income on Trip Purpose

But at the same time if commuters of age between 26 to 40 years considered, majority (83%) of commuters going for work trips belongs to monthly income of 5000 to 20,000. Now for the commuters having age above 40 years, 75% commuters have monthly income varies in between 10,000 to 50,000.

With the progression of age, person’s willingness to use the public transportation system is more than the person with same income but younger.

G. Trip Length

% trip distribution for various trip lengths is given below.

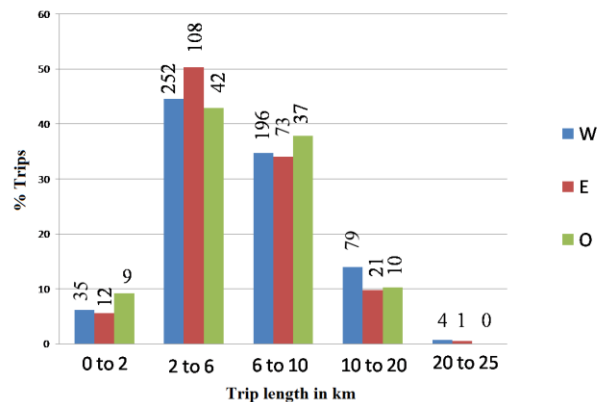


Fig. 7. Trip Length Distribution

Respectively 50%, 45% and 43% trips for education, work and other purpose are formed with trip length in a range of 2-6 km. Average trip length for all purpose is around 7 km.

The same trip length is validated with the secondary data source from competent authority. Data regarding the number of buses in operation, number of trips for each bus, total ticket fare collection during a day, total number of commuters using BRTS per day etc. are collected and analyzed for average trip length which shows the range in between 6 to 8 km.

**H. Compliments/Complaints for the BRTS**TABLE 2  
Compliments for BRTS

COMMENTS (COMPLIMENTS)	No. OF COMMENTS	%
Satisfied from Present Service	143	56.97%
Economic option for lower middle class	38	15.14%
Good for occasional trips	39	15.54%
Optimal option in era of inflation (mainly Fuel price)	31	12.35%
	251	100.00%

Objective of BRTS is satisfied from the point of view of 57% commuters which signify the usefulness of system. It is interesting to see that 12.35% commuters using BRTS due to rise in petrol price and 15.54% commuters using BRTS occasionally.

TABLE 3  
Complaints for BRTS

COMMENTS (COMPLAINTS)	No. OF COMMENTS	%
Delay	43	10.75%
Parking Problem at Bus station	49	12.25%
Overcrowded and less seating capacity of bus	128	32.00%
Operational problem at entry/exit	32	8.00%
Non optimized bus frequency	100	25.00%
Sudden jerks	26	6.50%
BRTS blocking traffic of other modes	22	5.50%
	400	100.00%

Complaints for the BRTS are high as compared to compliments by comparing the number of comments. 63.5% commuters are not happy with the present service level provided by authority. Service level defined in terms of comfort during the journey. Discomfort may attribute due to overcrowded bus, sudden jerks and most important is non-optimized bus frequency.

More than 12.5% commuters using their own vehicles to reach/leave from BRTS bus station as this community facing a problem of parking at BRTS bus station.

**IV. SUMMARY****A. Conclusions**

Shifting to BRTS from other than Public transport system (means other than AMTS) is around 57%. Quantification of benefits in terms of time saving, fuel saving, pollution control, reduction in accident rates, etc. due to BRTS should be investigated considering before–after comparison.

86% commuters belong to middle or lower middle class with monthly income less than 20,000. Also 73% work trips are from the commuters with monthly income in the range of

5000 to 20,000. Justification of dedicated BRTS corridor passing through normal traffic should be studied for time value of road users (Travelling on BRTS and other than BRTS).

**B. Future study scope**

Comments (complaints) of present commuters should be studied and corresponding measures should be taken to improve the reliability of public transportation system. All below listed studies are required for continuous supervision of BRTS which ultimately form a databank.

BRTS bus frequency should be studied in the context of daily commuter data between a pair of origin-destination. Optimization of frequency will increase the carrying capacity of route with comfort.

Sudden jerks at junction or station lead to the study of change in speed during a trip using GPS and driver should be trained to keep the movement of bus in possible uniform speed.

Entry-exit problem from BRTS bus to station is common. Use of same passage for entry-exit operation, creates conflict during entry-exit. Ideally entry–exit should be kept separate from each other and feasibility for the same can be studied.

**REFERENCES**

- [1] “Study of urban traffic pattern for mass transit system” B.Tech- Final Year Civil Engineering Project (2011), Institute of Technology, Nirma University.
- [2] Mukti Advani, Gitam Tiwari (2005) “Evaluation of public transport system – case study of Delhi metro,” in *START Conference held at IIT Kharagpur, India*.
- [3] Dr. Santosh AJALIHAI, Dr. T.S.Reddy, KAYITHA Ravindrer (2005) “Traffic characteristics of India,” Proceedings of the Eastern Asia Society for Transportation Studies, Vol 5, pp. 1009 – 1024.
- [4] “Study of traffic and transportation policies and strategies in urban areas in India” Final Report (2008), Ministry of Urban Development, Govt. of India.
- [5] “Service level benchmarks for urban transport,” Ministry of Urban Development, Govt. of India.
- [6] Venkatachalam THAMIZH ARASAN, Perumal VEDAGIRI (2011) “Modeling model shift from personal vehicles to bus on introduction of bus priority measures,” *Asian Transport Studies*, Vol -1, pp. 288 – 302.
- [7] “Performance report of BRTS – Ahmedabad (2009),” Center of Excellency in Urban Transport, CEPT University, Ahmedabad.
- [8] “Detailed project report – Ahmedabad BRTS (2005),” Gujarat Infrastructure Development Board. Govt. of Gujarat.



# Reduction In Co2 Emissions With Respect To Commuters' Shift From Private Transport To Proposed BRTS Corridor : A Case Study Of Patna.

A. Prof. Desai N A ,

B. Dr. Joshi G.J

C. Mr. Devendra Mokal

**Abstract**—Transportation demand has been substantially increased in most Indian cities especially in last two decades due to growth in commercial and industrial activities, urbanization, increase in household income, change in lifestyle etc. With increasing transport demand, together with increasing number of vehicles on road, problems associated to traffic congestion, road accidents and environmental pollution have amplified significantly. Transport sector is one of the major contributor to air pollution worldwide. According to Sharma and Mishra (1998), ambient air pollution in terms of Suspended Particulate Matter (SPM) in all metropolitan cities in India exceeds the limit set by World Health Organization (WHO) which leads to several health problems. Urban public bus transportation is the paramount for the growth and development of any city. However compared to major cities of world, statistics of Indian cities are showing very less percentage share of public transport almost in all major cities except few. This paper provides an overview on a proposed, Bus Rapid Transit ( BRT) corridor in a Patna city in the state of Bihar, India. Using revealed and stated preference data, probable mode shift to the new public transport mode ( BRT) has been evaluated and a binary logit model for travel mode choice have been developed. The model output pertaining to the number of commuters shifting from private mode to public transport modes has been used to calculate reduction in vehicular emissions due to shift of commuters from personalized to public transport mode.

**Index terms** - Urban transportation, Air pollution, Vehicular emissions, Mode choice, Modal shift.

## I. INTRODUCTION

Recently climate change has become a corporeal issue that must be addressed to avoid major environmental consequences in the future. Changes observed in public views has been due to the physical signs of climate change viz., rising sea levels, more severe storm and drought events, melting glaciers, rise in a global temperature, ozone layer depletion etc. There is general agreement among the world's climatologists that global warming is largely the result of greenhouse gas emissions from human activity (Pew Center 2006). More than 140 nations have signed the Kyoto Protocol, making

a commitment to reduce their greenhouse gas (GHG) emissions by 5.2 percent from 1990 levels by 2012. Various sectors are responsible for emissions of green house gas emission out of that transportation is a major contributor of carbon dioxide (CO<sub>2</sub>) and other greenhouse gas emissions, accounting for approximately 14 percent of total anthropogenic emissions globally. Emitted gasses from vehicles include carbon dioxide, methane, nitrous oxide, and other six kinds of greenhouse gases out of which carbon dioxide has the greatest impact on climate change.

## II. URBAN TRANSPORTATION SCENARIO IN INDIA AND WORLD

World experienced tremendous growth of population in last two decades. India is poised for rapid economic growth. It is projected that India's urban population would grow to about 473 million in 2021 and 820 million by 2051, as against only 285 million in 2001. Hence, satisfying transportation demand of current and future population is the most challenging job. Passenger mobility in urban India relies heavily on its roads, especially the bus transport will have to play a major role in providing passenger transport services, but easy availability of inexpensive fuel, the long-term favor of two wheeler & car as the predominant mode of urban travel by policymakers, and a limited public awareness concerning the negative health and environmental impacts of excessive auto use have aggravated existing travel related problems. With deteriorating levels of mass transport services and increasing use of personalized modes, vehicular emission has reached an alarming level in most Indian cities. Growing traffic and limited road space have reduced peak-hour speeds to 5 to 10 kms. per hour in the central areas of many major cities. This also leads to higher levels of vehicular emission. In India public transport services are inefficient and inadequate. Urban areas in India, which include a wide range of mega cities and metro cities are not all that fortunate in terms of public transportation. Transport in this context has been a victim of ignorance, neglect, and confusion. Each category of city is facing different types of problems related to transportation system. The metropolitan cities having more share of public transportation are lacking in the level of service. Therefore there is a need to evaluate entire transportation system from users view point before suggesting any solution. Efficient public transport system will definitely affect the low capacity personalized modes in terms of usage. Reductions in these modes will results into reducing effective road space which will reduce fuel and energy consumption leading to sustainability of the system. A high level of air and noise pollution, travel delay, road accidents etc. are the

A - Associate Professor, Civil Engineering department, Babaria Institute of Technology, Varnama, Vadoadra.

B - Associate professor, Civil Engineering department., S.V. National Institute of technology, Surat-395007

C - Mr. Devendra Mokal Former P.G, Student Civil Engineering department., S.V. National Institute of technology, Surat-395007

undesirable feature of overloaded road networks of urban area. The main reasons for these problems are the prevailing imbalance in modal split, inadequate transport infrastructure, and its suboptimal use. Public transport systems have not been able to keep pace with the rapid and substantial increases in demand over the past few decades. Bus services in particular have deteriorated, and their relative output has been further reduced as passengers have turned to personalized modes and intermediate public transport. There must be a general recognition that without public transport cities would be even less viable and policy should be designed in such a way that it reduces the need to travel by personalized modes and boosts public transport system. There is a need to encourage public transport instead of personal vehicles, this requires both an increase in quantity as well as quality of public transport and effective use of demand as well as supply-side management measures for healthy future.

#### IV. LITERATURE REVIEW

It is predicted that by the year 2050, global carbon emissions from transport will increase by 30% to 50% compared with the year 2008 (Huang Rui, 2008). The Oslo Climate and Environmental International Research Center published a report in the "National Academy of Sciences" magazine of the United States in 2008 which showed that emissions released from fuel used by automobiles, ships, airplanes, and trains was one of the main factors that contributed to global warming. In the developing countries, air quality crisis in cities is attributed to vehicular emission which contributes to 40-80% of total air pollution (Ghose et al., 2005). To reduce vehicular air pollution, many studies have been proposed in the past few years. Sterner et al. 1992, Morikawa and Sasaki 1993, Hayashi et al. 1995, Ito et al. 1995, Koopman 1995, Sano et al. 1995, Matsuoka et al. 1997, Ohno et al. 1997, Sugie et al. 1997 has given their valuable contribution for the analysis of reduction in pollution using various surveys and methods. According to Yoshida and Harata (1996) Shifting travellers' travel mode from the private vehicle to public transport is one effective method of reducing CO<sub>2</sub> emissions and easing traffic congestion. It is concluded by Philemon Kazimil Mzee, Yan Chen (2010) in their study that the introduction of a BRT system in Dar es Salaam has numerous evident gains in terms of improving the current public transport supply. Based on the French national traffic survey data in 1994, Jean-Pierre Nicolas and Damien David analyzed the carbon dioxide emissions of residents' daily travel and their impact factors (Jean- Pierre Nicolas, Damien David, 2009). From a comparative perspective of the city bus and car, Fengbao Wang discussed the strategies of energy saving and emission reduction of urban transport (Wang Fengbao, 2008). According to the development trends of motor vehicles in China and their fuel consumption trends in the next twenty years, Yali Yu summarized the application of renewable energy in vehicles, and put forward some suggestions on sustainable development of

vehicles (Yu Yali, 2007). Tian XIAO, Xiaofa SHI and Ke WANG (2010) has utilized two methods suggested in (GHG Protocol, 2003) for estimation of carbon dioxide emissions for the cities like Huizhou, Dongguan, and Rizhao.

#### III. IMPACT OF VEHICULAR EMISSIONS ON AMBIENT AIR QUALITY

Environmental pollution is a common problem in both developing and developed countries. Every year large quantities of toxic wastes are discharged into the environment from the ever increasing production of goods, transportation activities and from the burning of fossil fuels. Discharged gases from different type of vehicles include carbon dioxide, methane, nitrous oxide, and other six kinds of greenhouse gases. Carbon dioxide has the greatest impact on climate change. There are many sources of green house gas emissions but those generated by the use of transport account for a large proportion of total emissions. The urban population is mainly exposed to high levels of air pollution including metals from vehicular exhaust. These particles can penetrate deep into the respiratory system, and studies indicate that the smaller the particle, more severe the health impacts (Pope et al., 1995). Symptoms like chronic cough, wheezing and breathlessness have been reported on exposure to these pollutants (Chabra et al., 2001). India has 23 major cities of over 1 million people and ambient air pollution exceeds the WHO Standards in many of them. (Gupta et al., 2002.)

#### V. CASE STUDY : AN INTRODUCTION

Patna is the 5<sup>th</sup> fastest growing city in India and being the capital of Bihar the city has acquired a strong position in regional trade, cultural and business. It is also the centre for education and health facilities in the state. A software technology park is located in the city and more thrust for IT has been envisaged. The industries existing around include steel casting, cotton mills, warehousing, electronics, leather & shoes etc. More than 250 industries has been identified in and around city. The existing circulation pattern of Patna is of linear type and approximately 25 km long and 9 km to 10 km wide. According to the census 2011 present population of Patna is around 2.2 million which is expected to cross 2.8 million in the year 2021. Patna registered a 67 fold increase in number of vehicles in two decades (1981-2001). In 1981 the number registered motor vehicles were 4384 and it increased to 2,94,164 in 2001. From 1996 to 2001, vehicular population in Patna grew at an average annual rate of around 6%, whereas for period from 1981 to 2001, annual growth rate was 23%. In the registered vehicles the total number of cars and two wheelers adds up to around 75-89%. Looking to the recent trend towards sudden increase in personalized vehicle usage and rise in air pollution level in the city, Patna has been selected for developing a model of mode shift behavior and probable reduction in air pollution with introduction to Bus Rapid Transit.

### A Study Objectives

1. To estimate rider ship for transit systems through likely mode shift from existing mode to the proposed bus rapid transit system.
2. To estimate reduction in CO<sub>2</sub> emissions due to probable shift of personalized and public bus commuters to rapid bus transit system.

### B Need And Significance Of The Study

In last few years, the ownership and usage of private vehicles in Patna has increased phenomenally. Research in similar situation elsewhere has shown that if there is corrective actions, than the sustainability of the urban transport system will be in a very critical situation. As a matter of fact the transportation policies and management almost in all Indian cities has not strongly taken side to support increased usage of public transport and decreased usage of personalized vehicles. Deterioration in quality of public service has caused negative shift, which is a frightening situation for urban governments. The study has also chosen with an initiative to predict reduction in air pollution by obtaining modal shift from other travel mode to environment friendly rapid bus transit system. The findings from the study can be useful to government bodies and decision making authorities to devise long term sustainable transportation plans and policies as well as short term implementation plans.

### C Data Collection

As per division of election wards in Patna municipal corporation, total 72 zones are selected for conducting house hold interviews using random sampling method. A survey questionnaire is designed and data has been collected through stated and revealed preference techniques on the selected corridor of Bus Rapid Transit in Patna city. Socioeconomic characteristics, travel characteristics and mode choice preferences for hypothetical route have been included in the surveys. In addition to this to fulfill data gaps necessary information have been collected from different agencies viz. municipal corporation, RTA office, urban development authority, traffic cells etc.

### D Identification Of Proposed Brts Corridor

Inventory of road network for selection of BRTS corridor is carried out. Availability of road width for lane of BRT system is examined and according to that corridor had been finalized for proposed BRTS corridor. Total length of the BRTS corridor is 19.6 Km, in which major junctions like Saguna more, Baily road, Raza Bazar, Khojpora, IAS colony, Patna junction,

Kankarbaug, Chitagupta nagar, Meethapur, Rajendra nagar railway station, Agam kaur, Kumhar and Gahdhi setu bridge which covers almost 35% of city traffic.

### E Methodology

The method employed in the study was clarification of the problem, statement of objectives, data requirement, data collection and processing, analysis of the data, and finally, the evaluation of results. Home Interview survey has been conducted to obtain socio-economic information, travel characteristics and required information to estimate probable modal shift. To understand Socio-Economic background three parameter of the family are taken into consideration viz. vehicle owner-ship, house-hold size and income-group distribution. Attributes influencing mode choice of commuters are identified and relevant literature is referred. Questionnaire has been designed and data have been collected on the selected proposed BRT route. Data is analyzed using disaggregate approach for household and travel characteristics. A discrete choice model predicts a decision made by an individual as a function of any number of variables. These models consider that the demand is the result of several decisions of each individual in the population under consideration. These decisions usually consist of a choice made among a finite set of alternatives. The model can be used to estimate the total number of people who change their behavior in response to an action. As a result, the change in choice of travel mode can be estimated. Willingness to shift data is analyzed and commuters travel time and travel cost is taken in to account from it. The income group wise binary logit models are developed using GRETl software. Probable shift to BRT is estimated under various operating scenarios of the proposed transit system. Finally from the average estimated modal shift reduction in air pollution is calculated.

### F Data Analysis And Travel Mode Shift Calculations

Utility function for travel mode choice developed in GRETl platform has been used to develop the mode shift models for all income groups. The same utility function is used to calculate the shift from existing mode to proposed BRT using binary Logit models. For every income group probabilities of shift are estimated for all available modes keeping travel time and travel cost as variable. Figure – 1 shows existing modal split. The results obtained from model regarding modal shift are presented below in Table – 1. Note : Shift from car users have been almost negligible and hence not considered in final calculations which shows that mostly all car users can be considered as captive riders of the mode.

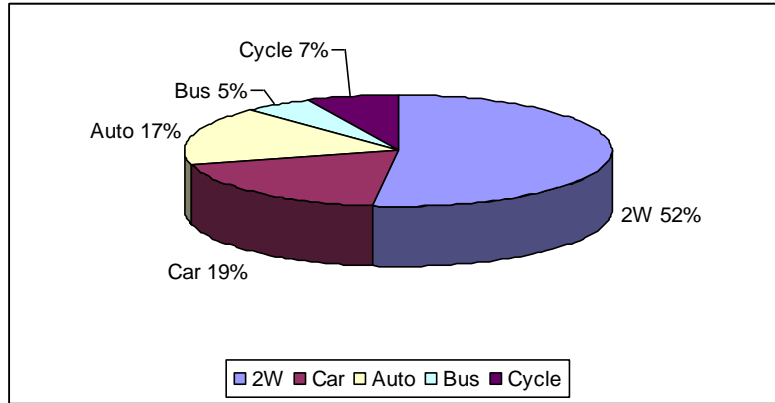


Fig -1 Existing modal split in Patna city.

G Reductions In Co<sub>2</sub> Emissions Due To Travel Mode Shift  
Based on the results obtained from Logit model

percentage shift to BRTS has been calculated and results are presented in following Table – I

TABLE – I  
MODAL SHIFT CALCULATIONS BASED ON THE RESULTS OF LOGIT MODEL.

Sr.No	Income group	Trip length ( Km)	% shift to BRT from other modes			Person shift to BRT from other modes		
			2W	Auto	Bus	2W	Auto	Bus
1	LIG	00-05	2.40	5.10	95.70	7862	5462	30146
		05-10	1.30	8.40	95.10	4259	8996	29957
		10-15	0.80	18.40	92.10	2621	19706	29012
		15-20	0.80	40.80	86.30	2621	43697	27185
2	MIG	00-05	2.10	4.60	3.20	6880	4927	1008
		05-10	1.10	7.70	3.30	3604	8247	1040
		10-15	0.80	18.10	4.80	2621	19385	1512
		15-20	0.80	42.70	9.00	2621	45732	2835
3	HIG	00-05	2.10	4.60	3.40	6880	4927	1071
		05-10	1.30	7.70	4.40	4259	8247	1386
		10-15	1.30	18.20	8.70	4259	19492	2741
		15-20	1.60	42.70	21.40	5242	45732	6741
<b>Total Shift mode Wise</b>						53726.4	234549	134631
Note : Population of Patna city taken as reference for calculatio = 1.8 million based on 2001 census.Out of which 35% population is corridor specific i.e 0.63 million. Person shift is calculated based on existing modal split and percentage shift obtained from model with reference to total population.								

H Selection Of Emission Factors And Calculation For Reduction In Co<sub>2</sub> Emission.

Currently, there is no uniform standard with regard to emission factors for carbon dioxide. Models presently used are all either too complicated for calculations of carbon emissions of urban passenger transport or requires huge number of complex data. Therefore, the emission factors this paper use come from the report “Traffic of China: the energy consumptions and emissions of different transportation modes”, which was finished by Institute for Energy and Environmental

Research Heidelberg & Comprehensive Transportation Institute of Chinese National Development and Reform Commission. The emission factor for each transportation mode is listed and reduction in CO<sub>2</sub> emissions has been calculated which is represented in following Table – II.

Sr.No	Discription	Existing scenario			Scenario after implementation of BRTS
		2W	Auto	Bus	
1	Total person travel mode Wise, which are expected to shift to BRTS based on developed model.	53726	234549	134631	422906
2	Emission factor for each transport mode in gm/km. person	82	120	185	88
3	Emission due to usage of different transport mode in gm/km	4405565	28145880	24906735	37215763
4	Emission due to usage of different transport mode in tones/km	4.41	28.15	24.91	37.22
Total CO <sub>2</sub> emissions in tones / km		57.46			37.22
Total reduction in CO <sub>2</sub> emissions in tones / km					20.24

Source :Emission factor used for calculation has been finished by Institute for Energy and Environmental Research Heidelberg & Comprehensive Transportation Institute of Chinese National Development and Reform Commission.

## VI. CONCLUSION

Due to rapid economic growth and urbanization, Indian cities are facing critical position in satisfying the transportation demand of current and future population. Bus transport will have to play a major role in providing passenger transport services, but easy availability of inexpensive fuel, long-term favor of two wheeler & car as the predominant mode of urban travel and unawareness at commuters end about critical environmental impact have aggravated existing travel related problems. Even though public transport is having a very good potential to

resolve congestion and pollution related issues it is observed that the system has been a victim of ignorance, negligence and confusion. From the study it is concluded that if public transport services are improved as per user requirement and reduction in travel time and cost is implemented through various actions, commuters are willing to shift from other travel modes to proposed BRTS. From the model results it has been obtained that over all shift from 2W, Auto and existing buses is almost 0.42 million commuters/day to proposed BRTS corridor, which leads to reduction in CO<sub>2</sub> emissions as 20.24 tones/km.

## REFERENCES

- [1] Air pollution with special reference to vehicular pollution in urban cities, an online report by India State of the environment: 2001.
- [2] Atsushi Fukuda, Tuenjai Fukuda, Pongrid Klungboonkrong, Yasuki Shirakawa, Komei Yamaguchi and Atit Tippichai, " Study on Impacts of Mass Rapid Transit Development in Developing Cities - A case study of the subway of the extension blue line in Bangkok, Thailand ", online notes.
- [3] H. TAMURA (2003), "Estimation of a reduction in CO<sub>2</sub> emissions by shifting commuters' travel mode from the private car to public transport" published in International Journal of Systems Science volume 34, number 3, 20 February 2003, pages 159-165.
- [4] Devendra Mokal & Dr. G.Joshi ( 2010), " Disaggregate mode shift behaviour study with respect to bus rapid transit system : Case study of Patna., dissertation submitted to SVNIT, surat for award of mtech degree.
- [5] Juanjuan LIU ( 2010 ), " Study on passenger car equivalents based on the emissions. "
- [6] Mario Cools, Elke Moons, Brecht Janssens & Geert Wets ( 2009), " Shifting towards environment-friendly modes: profiling travelers using Q-methodology", Springer Science+Business Media, LLC. 2009, Transportation (2009) 36:437–453.
- [7] Masami Kojima, Carter Brandon & Jitendra Shah ( 2000 ), " Improving Urban Air Quality in South Asia by Reducing Emissions from Two-Stroke Engine Vehicles", a world bank report on Traffic images of South Asia.
- [8] N Sharma, S Gangopadhyay and R Dhyani,(2010) "Methodology for estimation of CO<sub>2</sub> reduction from mass rapid transit system (MRTS) projects" Journal of scientific and industrial research Vol. 69, August 2010, pp. 586-593.
- [9] P. Koushki, P. ASCE, S. Al-Fadhala, O. Al-Saleh3 and A. H. Aljassar ( 2002), "Urban Air Pollution Impact of Modal Shift in School Transportation in Kuwait", published in Journal of urban planning and development , June 2002.
- [10] Philemon Kazimil ,Yan Chen ( 2010 ) , "Implementation of BRT System in Developing Countries as the Best Option in Reducing Emission the Case Study of Dart System in Dar es Salaam", IEEE 532.
- [11] P. Sharma, and N Mishra (1998), " Transportation for healthy tomorrow, issues and options, presented during the seminar on planning Delhi : Healthy city in the next millennium, DRC, ITPI, New Delhi.
- [12] Tian XIAO, Xiaofa SHI and Ke WANG ( 2010 ), "Analysis on Influencing Factors for Carbon Emissions of Urban Passenger

- Transport”, ICCTP 2010: Integrated Transportation Systems, ASCE.
- [13] Urban Air Pollution Control an online report by central pollution control board, govt. of India.
- [14] Vineeta Shukla, Poonam Dalal and Dhruva Chaudhry ( 2010 ) “Impact of vehicular exhaust on ambient air quality of Rohtak city, India” Journal of Environmental Biology November 2010, 31(6) 929-932.
- [15] William Vincent, Lisa Callaghan Jerram ( 2006 ), “The Potential for Bus Rapid Transit to Reduce Transportation-Related CO2 Emissions “ Journal of public transportation, 2006 BRT special addition.
- [16] Zhao Hongyu & Chu Yun ( 2009), " Study on Urban Design Strategy for Low-carbon Trip", published in the proceedings of International Conference on Energy and Environment Technology, IEEE.









# Workability of Self-Compacting Concrete

A. M.A.Farediwala , B. M.A.Jamnu

## **Abstract:**

Self-Compacting Concrete is a fluid mixture suitable for placing in structure with congested Reinforcement without vibration. This paper presents the result of an experimental research on the workability of Self-Compacting Concrete. The concrete mixes contained three different dosages of the hypo super plasticizer with fly ash. The percentage of fly ash that replaced cement in this research was 15%. The workability test utilized in this research were the Slump flow, V-funnel and L-box, which can be used to evaluate the passing ability of Self-Compacting Concrete.

## **Introduction:**

Self-compacting concrete (SCC), developed first in Japan in the late 1980s, represents one of the most significant advances in concrete technology in the last two decades. SCC was developed to ensure adequate compaction through self-consolidation and facilitate placement of concrete in structures with congested reinforcement and in restricted areas. SCC can be described as a high performance material which flows under its own weight without requiring vibrators to achieve consolidation by complete filling of the formworks even when access is hindered by narrow gaps between reinforcement bars.

Modern concrete production systems have not eliminated the need to monitor concrete Workability in the field. To the contrary, the advent of new so-called high-performance concrete mixes that are susceptible to small changes in mix proportions has made monitoring workability even more critical. Indeed, a National Ready-Mixed Concrete Association survey identified the Need for a better method to characterize the workability of high performance concrete (Ferraris And Lobo 1998). After more than 80 years of efforts, the concrete industry is still faced with the Challenge of developing a field test to measure the relevant rheological properties of concrete quickly and accurately.

Self-compacting concrete (SCC) is reduce labour in the placement of concrete by eliminating and reducing the need of vibration to achieve consolidation. Therefore, the main property that defines SCC is high workability in attaining consolidation and specified hardened properties. A highly flowable concretes not necessarily self compacting because SCC should not only flow under its own weight but should also fill the entire form and achieve uniform consolidation without segregation. One type of SCC is used in structure with closely spaced reinforcing bars and should be able to flow through completely fill the form without vibration. This characteristic of SCC is called the filling capacity.

A. Student of M.E.(CASAD),Department of Applied Mechanics  
Government Engineering College Dahod-389151 Gujarat , India  
B. Professor , Department of Applied Mechanics Government Engineering  
College Dahod-389151 Gujarat , India

SCC requires a reliable control of materials characteristics, mixing process and transformation. To obtain SCC mixture, the paste content (including mineral additions and super plasticizer dosage) has to be increased and the coarse aggregate content must be reduced. The “excess paste” should be of minimal quantity to create a lubricating layer around the aggregate particles and to achieve self compatibility. The paste in SCC behaves as the vehicle for the transportation of the aggregates. Minimum paste required in case of SCC is of interest for both economical and material performance reasons. Paste volumes are usually higher than convection. The purpose of this concrete concept is to decrease the risk due to the human factor, to enable the economic efficiency, more freedom to designers and constructors and more human work.

With the preface of the new generation of superplasticizers, SCC has been industrialized. This type of concrete having advanced viscosity and workability properties can easily fill the mould without the necessity of using vibrators. It is apparent that workability depends on a number of interacting factors such as water content aggregate type and grading, aggregate to cement ratio, kind and dosages of superplasticizers, and the fineness of cement. The main factors on self-compacting concrete are the water and superplasticizer contents of the mix since by simply adding them the interparticale

lubrication is increased.

The strict definition of workability is the amount of useful internal work necessary to produce full compaction. The useful internal work is a physical property of concrete and is the work or energy required to overcome the internal friction between the individual particles of the mixture. Because of the very high workability of self-compacting concrete, it needs no external vibration and can spread into place, fill the framework and encapsulate reinforcement without any bleeding or segregation. In other words, to ensure that reinforcement can be encapsulated and that the framework can be filled completely, a favorable workability is essential for self-compacting concrete. Moreover, aggregate particles in self-compacting concrete are required to have uniform distribution in the specimen and the minimum segregation risk should be maintained during the process of transportation and placement.

## **Material and Mix proportion:**

This part of the paper presents the specifications of the mixes used for obtaining the workability of self-compacting concrete. The cementations materials used were pozzolana Portland cement and fly ash, Natural river sand crushed aggregates with 20mm, The control mixes were prepared by replacing 15% of the cement with fly ash on mass-for- mass basis. The details of the mix proportions of the present research are given in Table. As a result of different dosages of the super plasticizer, In this research used the super plasticizer is a hypo super plasticizer for durable and heavy duty concrete (As per

IS 9103-1999).it is based on blends of organic polymers and is of chloride free& nontoxic.

**Test Method:**

The ingredients for the various mixes are weighed and mixing is carried out using rotator type concrete mixer. Precautions are taken to ensure uniform mixing of ingredients. The batching sequence consisted of homogenizing the sand and coarse aggregate for 30 sec, then adding some amount of the mixing water (water+superplasticizer) into the mixer and continuing to mix. Then Remaining water is added into the mixer and continued rotation till the SCC is achieved. After getting SCC slump flow, V-funnel and L-box tests are performed to get required responses for fresh concrete properties.

**Slump Flow Test:**

The slump-flow was a test to assess the flowability and the flow rate of self-compacting concrete in the absence of obstructions. The result was an indication of the filling ability of self-compacting concrete.

It is the most commonly used test, and give a good assessment of filling ability. It gives no indication of the ability of the concrete to pass between reinforcement without blocking, but may give some indication of resistance to segregation. The fresh concrete would poured into a cone as used for the slump test. The largest diameter of the flow spread of the concrete and the diameter of the spread at right angles to it are then measured and the mean is the slump-flow.



Fig 1 Slump flow

**V-Funnel Test:**

The equipment consists of a V-shaped funnel shown in fig An alternative type of V-funnel, the O funnel, with a circular section is also used in Japan. The V-funnel test used to assess the viscosity and filling ability of self-compacting concrete. A V shaped funnel is filled with fresh concrete and the time taken for the concrete to flow out of the funnel is measured and recorded as the V-funnel flow time.The described V-funnel test is used to determine the filling ability of concrete with a maximum aggregate size of 20 mm.

Mix components	Concrete mixes		
	SCC1	SCC2	SCC3
Cement(kg/m <sup>3</sup> )	425	446	468
Fly ash(kg/m <sup>3</sup> )	75	79	83
Aggregate(kg/m <sup>3</sup> )	870	850	800
Snad(kg/m <sup>3</sup> )	670	850	850
Super plasticizer (ml/bag)	500 to 700	500 to 700	500 to 700



Fig 2 V-funnel

**L-Box Test:**

The L-box test is used to assess the passing ability of self-compacting concrete to flow through tight openings including spaces between reinforcing bars and other obstructions without segregation or blocking. There are two variations; the two bar test and the three bar test. The three bar test simulates more congested reinforcement. A measured volume of fresh concrete is allowed to flow horizontally through the gaps between vertical, smooth reinforcing bars and the height of the concrete beyond the reinforcement is measured.

The passing ratio *PL* or blocking ratio *BL* is calculated using equation

$$\begin{aligned}
 PL &= H/H_2 \\
 \text{Or} \\
 BL &= 1-H/H_2 \\
 \text{Where } H &= H_1 - \Delta h
 \end{aligned}$$



Fig 3 L-Box

**Workability of Self-Compacting Concrete**

Concrete mixes		Superplasticizer dosage (ml/bag)	Workability Tests		
			Slump flow (mm)	V-funnel (second)	L-box (ratio)
SCC1	0.40	500	640	9.0	0.79
		600	675	8.5	0.81
		700	690	8.3	0.82
	0.50	500	725	7.5	0.90
		600	760	6.0	0.93
		700	790	5.4	0.95
SCC2	0.40	500	630	9.0	0.88
		600	670	8.7	0.89
		700	710	8.1	0.85
	0.50	500	720	7.2	0.93
		600	740	6.7	0.95
		700	755	6.1	0.98
SCC3	0.40	500	660	10.1	0.8
		600	690	9.7	0.83
		700	710	9.2	0.86
	0.50	500	700	8.0	0.92
		600	730	7.6	0.96
		700	755	6.5	0.97

**Conclusions:**

From the results presented in this paper, using containing different dosages of a kind of superplasticizer, the main conclusions are:

- (1) Overall cost of SCC may be slightly higher compared to same grade of conventional concrete mainly due to cost of chemical admixtures. However, partly the additional cost is compensated by reducing the cost of compaction and achieving speed during construction.
- (2) SCC gives good finishing as compared to ordinary concrete without any external mean of compaction.
- (3) The effect of fly ash and the dosage of the superplasticizer were higher on improving the workability.

**REFERENCES:**

- [1] M.A.Ahmadi, O. Alidoust, I.Sadinejad, and M.Nayeri "Development of Mechanical Properties of Self-Compacting Concrete Contain Rice Husk Ash" World Academy of Science Engineering and Technology, 34, 2007
- [2] Zoran Grdic, Iva Depotovic, Gordana Toplicic Curcic "Properties of Self-Compacting Concrete with Different Types of Additives", Architecture and Civil Engineering, Vol.6, N<sup>o</sup>2, pp.173-177, 2008.
- [3] Dr.R.Sri Ravindrarajah, F.Farrokzadi and A.Lahoud "Properties of Flowing concrete and self-compacting concrete with high Performance Superplasticizers" RILEM Publications, Vol.1, pp.1048, August 2003.
- [4] C.Selvamony, M.S.Ravikumar, S.U.Kannan, and S.Basil Ganapappa "Investigation on self-compacted Self-curing concrete using Limestone Powder and Clinkers." APRN Journal of Engineering and Applied Sciences Vol.5 No.3, March 2010.
- [5] Cristian Druta "Tensile strength and bonding characteristics of self-compacting concrete" A Thesis, August 2003.
- [6] Hajima Okamura and Masahiro Ouchin "Self-Compacting Concrete" Journal of Advance Concrete Technology, Vol.1, No.1, pp 5-15, April 2003.
- [7] Khayat, K.H., Assaad, J. and Daczko J. "Comparison of field-oriented test methods to assess dynamic stability of self-consolidating concrete," ACI Materials Journal, 101, 2, 2004 168-176.
- [8] Ding, Y.N., Liu, S.G., Zhang, Y. and Thomas, A. "The investigation on the workability of fibre Cocktail reinforced self-compacting high performance concrete," Construction and Building Materials, 22, 2008, 1462- 1470.
- [9] European Project Group, Specification and guidelines for self-compacting concrete, United Kingdom, EFNARC, 2002.
- [10] N.R.Gaywala, D.B.Raijiwala "Self-Compacting Concrete: A Concrete of next decade" Journal of Engineering Research and Studies E-ISSN0976-7916.

# Semi Active Control of Buildings Using Magneto-rheological Damper

A. I H. Vadatala, B. D P. Soni, C. K N. Sheth, D. V A. Shah

**Abstract**—In this paper, a six storey shear frame is numerically evaluated for semi-active control strategy using magnetorheological damper placed at ground and first floor. Clipped optimal control law with linear quadratic regulator is utilized to control the damper forces. Simulation of the closed loop force feedback system is done using SIMULINK, and state space approach is used to solve the dynamic equations, both for uncontrolled and controlled cases. The structural responses; displacement, inter-storey drift, and absolute acceleration for all stores at deck level for controlled and uncontrolled cases are evaluated and compared for three past ground motion history. The results showed that the semi-active control strategy using clipped optimal control law and linear quadratic regulator yields the best results as compared with the other control algorithms. The controller produces optimum forces and response reduction is far better than passive off and passive on cases.

**Index Terms**—Bouc-Wen, Clipped Optimal Controller, Magnetorheological damper, Semi-active control.

## I. INTRODUCTION

IN traditional approach, civil engineering structures resist vertical and lateral loads according to their stiffness which are dependent on member sizes. Larger dimensions are required to resist larger forces, widely known as demand and capacity based design. Vertical loads such as dead and live loads can be applied to greater accuracy and vary a little with the calculated. On the contrary the magnitude of lateral loads, especially seismic forces applied to analyze and design the structure can vary to great extent. Also it is not at all economical to design the structure to resist the most catastrophic earthquake forces. Hence a need to develop a more efficient and smarter way to reduce the structural response was highly expected, such as structural control engineering.

Earthquake forces impart energy to the structure, which produces push-pull effects to the structures causing displacements which in turn produce forces on the structure.

Preliminary, response of any structure to earthquake loadings is a function of its three main inherent property, namely mass, stiffness and damping. The primary focus was on altering, adding or modifying these three properties to reduce the structural response and control the structure. Since initial conceptual studies by Yao in 1972, structural control in civil engineering has gain attention and interest resulting in research and development in the field.

Structural control strategies were classified in four groups, namely passive, active, hybrid and semi-active control systems. Until about 1990, only passive control method was implemented for practical engineering systems which dissipate the earthquake energy to reduce the response. Due to limitations of passive control to become adaptable active control methods were implemented which overruled this limitation. Active control method utilizes actuators to both add and dissipate the energy. Semi-active control system combines the features of both, passive and active control strategies to reduce the responses under various earthquakes. They do not impart energy to the systems and requires very less power. Thus they result bounded-input and bounded-output.

In this study, clipped optimal controller using linear quadratic regulator is utilized for control forces produced by MR damper. Also simple Bouc-Wen model is used to represent the dynamics of MR damper for shear mode. A benchmark problem of Laura M. Jansen and Shirley J. Dyke is used for the study. The simulation for uncontrolled and controlled is developed in SIMULINK tool and the numerical example is tested for various ground history and the outputs are evaluated and compared.

## II. SEMI-ACTIVE CONTROL USING MR DAMPER

Semi-active devices are different from active devices in that semi-active devices can only produce dissipative forces. Semi-active devices include variable orifice dampers, variable friction dampers, controllable tuned liquid dampers, controllable fluid dampers, and etc these devices can be viewed as controllable passive devices, in that the characteristics of the passive devices can be changed in real time. In this manner, semi-active devices can produce the desired dissipative control forces.

MR fluids consist of micron-sized, magnetically polarizable particles dispersed in a liquid medium such as mineral or silicone oil. MR fluids are smart, synthetic fluids changing their viscosity from liquid to semi-solid state within milliseconds if a sufficiently strong magnetic field is applied. MR damper exhibits a variable damping coefficient depending

A. Mtech student at DDU, Nadiad, INDIA. Phone: 09824536846, E-mail: [ivadtala@gmail.com](mailto:ivadtala@gmail.com).

B. Associate Professor of Civil Department, SVIT, Vasad, INDIA E-mail: [soni\\_svit@yahoo.com](mailto:soni_svit@yahoo.com).

C. Head of Civil Department, DDU, Nadiad, INDIA. E-mail: [knsheth123@gmail.com](mailto:knsheth123@gmail.com)

V Shah, Professor Instrumentation & control Department, DDU, Nadiad, INDIA.

on the strength of an accompanying magnetic field. A high magnetic field creates a nearly unyielding damper filled with a semi-solid fluid while no magnetic field produces an ordinary viscous damper. MR fluid devices have been shown to mesh well with application demands and constraints to offer an attractive means of protecting civil infrastructure systems against severe earthquake and wind loading. Figure 1 shows schematic of 20t MR damper.

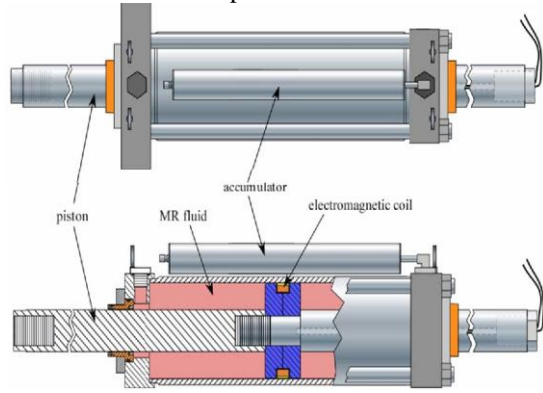


FIG. 1 Schematic of 20t MR Damper

In contrast to conventional electro-mechanical solutions, MR technology offers:

- Real-time, continuously variable control of Damping, Motion and position control.
- High dissipative force independent of velocity.
- High yield stress of about 100 kpa.
- Simple design as there are few or no moving parts.
- Quick response time of just 10 milliseconds.
- Consistent efficacy across extreme temperature variations range of -40 C to 150 C.
- Minimal power usage typically 12V, 1 Amp max current; fail-safe to battery backup, which can fail-safe to passive damping mode.
- Inherent system stability as no active forces are generated.
- MR fluids can be operated directly from low-voltage power supplies.
- Insensitive to impurities.

### III. MATHEMATICAL MODEL

Due to the inherent nonlinear nature of magnetorheological dampers, one of the challenging aspects for developing and utilizing these devices to achieve high performance is the development of models that can accurately describe their unique characteristics. In the past two decades, the models for MR dampers have been focused on how to improve the modeling accuracy. Although the force–displacement behavior is well represented by most of the proposed dynamic models for MR dampers, no simple parametric models with high accuracy for MR dampers is developed till date. In order to In order to model the behavior of the shear mode MR damper, Jansen and Dyke proposed a Bouc–Wen hysteresis operator based dynamic model for shear mode damper, as shown in figure 2.

The Bouc-Wen hysteretic behavior was initially formulated by Bouc as an analytical description of a smooth hysteretic model and later generalized by Wen; the hysteresis model of Bouc as modified by Wen, possesses an appealing mathematical simplicity and is able to represent a large class of hysteretic behavior, from inelastic stress–strain relationships found in structures to MR behavior. The Bouc-Wen model has been extensively applied to simulate the hysteresis loops since it possesses the force–displacement and force–velocity behavior of the MR dampers.

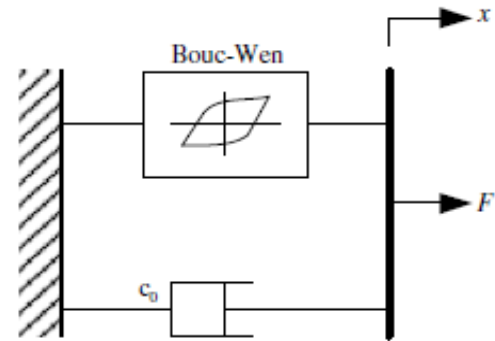


FIG. 2 Mechanical Model of the Parallel-Plate MR Damper

The governing equations of force produced by this model are as follows,

$$f = c_0 \dot{x} + \alpha z \tag{1}$$

$$\dot{z} = -\gamma |\dot{x}| z |z|^{n-1} - \beta \dot{x} |z|^n + A \dot{x} \tag{2}$$

$$c_0 = c_0(u) = c_{0a} + c_{0b} u \tag{3}$$

$$\alpha = \alpha(u) = \alpha_a + \alpha_b u \tag{4}$$

$$\dot{u} = -\eta (u - v) \tag{5}$$

Where,

$z$  is a hysteresis component that represents a function of the time history of the displacement.

$A, \beta, \gamma$  = shape parameters for hysteresis.

$n$  = controls smoothness from pre to post yield region.

$\alpha$  = scaling value for Bouc-Wen model.

$\eta$  = response time.

$C_0$  = viscous damping.

### IV. CONTROL ALGORITHM

One of the algorithms that have found to be very effective with the use of MR damper is clipped optimal controller, proposed by Dyke, et al. The clipped optimal design approach is to design a linear optimal controller ‘K’ that calculates a vector of desired control forces  $f = \{f_1, f_2$

...fn} <sup>T</sup>, based on measured structural response and applied force on the structure.

To induce MR damper to generate approximately the corresponding desired optimal control force  $f_{ci}$ , the command signal is  $V_i$  is selected as follows. When the  $i^{th}$  mode MR damper is providing the desired force, the voltage applied to the damper should remain at present value. If magnitude of force produced by MR damper is smaller than the magnitude of desired optimal force and the two forces have same sign, the voltage applied to the current driver is increased to maximum level so as to increase the force produced by MR damper to match the desired control force, else the command voltage is set to zero. Eq.6 correlates the voltage force relationship and the same can be graphically represented as shown in figure 3.

$$V = V_{max} H(\{F_{Ci} - Fi\} Fi) \quad (6)$$

$$f_{ci} \geq f_i, V_i = V_{max} \quad (7)$$

$$f_{ci} \leq f_i, V_i = 0 \quad (8)$$

Where,

H is the Heaviside step function or unit step function.

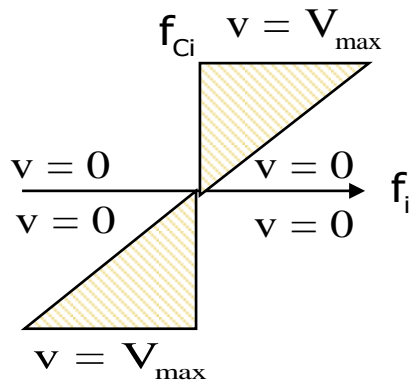


FIG. 3 Graphical Representation of Clipped Optimal Controller

Although variety of approaches may be used to design the optimal controller, Linear quadratic regulator (LQR) has been advocated for use, as it gives better results compared to its substitute, namely linear quadratic Gaussian.

#### V. LINEAR QUADRATIC REGULATOR

The design of a control system is an attempt to meet a set of specifications which define the overall performance of the system in terms of certain measurable quantities. This process becomes merely a trial and error process. If however a single performance index could be established on the basis of which one may describe the goodness of system response, then the procedure would become logical and straight forward. Thus a cost function or performance index is defined as in equation 9, using the Integral square technique, which will design a state variable feedback gain that is optimal.

$$J = \int_0^t [x^T Q x + u^T R u] dt \quad (9)$$

The objective of optimal state variable feedback design is to select the state variable feedback gain K which stabilizes the system and also minimizes the performance index J. Since the plant is linear and the performance index is quadratic, the problem of determining the state variable feedback which minimizes J is called the Linear Quadratic Regulator (LQR). In this case the desired performance of the system is given by the values of the two matrices Q and R present in the performance index to be minimized. The performance index J can be interpreted as an energy function, such that making it small keeps small the total energy of the closed-loop system. Note that both the state  $x(t)$  and the control input  $u(t)$  are weighted in J, so that if J is small, then neither  $x(t)$  nor  $u(t)$  can be too large. This guarantees that the closed-loop system will be stable.

The solution of equation 9, will find the optimal control force for the system. Hence to minimize J, such that

$$u = -kx \quad (10)$$

Where,

$$K = - R^{-1} B^T P \quad (11)$$

and

$$A^T P + PA + Q - PBR^{-1}B^T P = 0 \quad (12)$$

Equation 12 is known as Ricatti's equation, solving which yields P.

#### VI. SELECTING Q AND R MATRICES FOR PERFORMANCE INDEX

This performance index describes a balance between the amount of energy used to achieve desired performances and the specified desired performances. This balance is described by the choice of the two matrices Q (an 2n x 2n matrix) and R (an m x m matrix) which are selected by the design engineer. Depending on how these design parameters are selected, the closed-loop system will exhibit a different response.

- Selecting Q large means that, in order to keep J small, the state  $x(t)$  must be smaller. This means that larger values of Q generally result in the poles of the closed-loop system matrix  $A_c = (A - BK)$  being further left in the s-plane so that the state decays faster to zero.
- Selecting R large means that the control input  $u(t)$  must be smaller to keep J small. Larger R means that less control effort is used, so that the poles are generally slower, resulting in larger values of the state  $x(t)$ .
- One should select Q to be positive semi-definite and R to be positive definite. This means that - the scalar quantity  $x^T Q x$  is always positive or zero at each time t for all functions  $x(t)$ , and the scalar quantity  $u^T R u$  is always positive at each

time  $t$  for all values of  $u(t)$ . In terms of eigenvalues, the eigenvalues of  $Q$  should be non-negative, while those of  $R$  should be positive. If both matrices are selected diagonal, then all the entries of  $R$  must be positive while those of  $Q$  should be positive, with possibly some zeros on its diagonal. Note that if  $R$  is chosen positive definite then  $R$  is invertible.

VII. NUMERICAL EXAMPLE

To evaluate these algorithms for use with the MR damper, a numerical example is considered in which a model of a six story building is controlled with four MR dampers. Two devices are rigidly connected between the ground and the first floor, and two devices are rigidly connected between the first and second floors, as shown in Fig. 4. The maximum voltage input to the MR devices is  $V_{max} = 5$  V. The governing equations can be written in the form of (13) by defining the mass of each floor  $m$  as  $22.7$  N/(m/s<sup>2</sup>) the stiffness of each floor  $k$  as  $29700$  N/m and a damping ratio for each mode of  $0.5\%$ .

The MR damper parameters used in this study are,

$$\begin{aligned} \alpha_a &= 8.66 \times 10^2 \text{ N/m} & A &= 120 \\ \alpha_b &= 8.66 \times 10^2 \text{ N/m} & \gamma &= 300 \times 10^4 \text{ m}^{-1} \\ C_{0a} &= 0.0064 \text{ N.sec/m} & \beta &= 300 \times 10^4 \text{ m}^{-1} \\ C_{0b} &= 0.0064 \text{ N.sec/m} & \eta &= 80 \text{ sec}^{-1} \end{aligned}$$

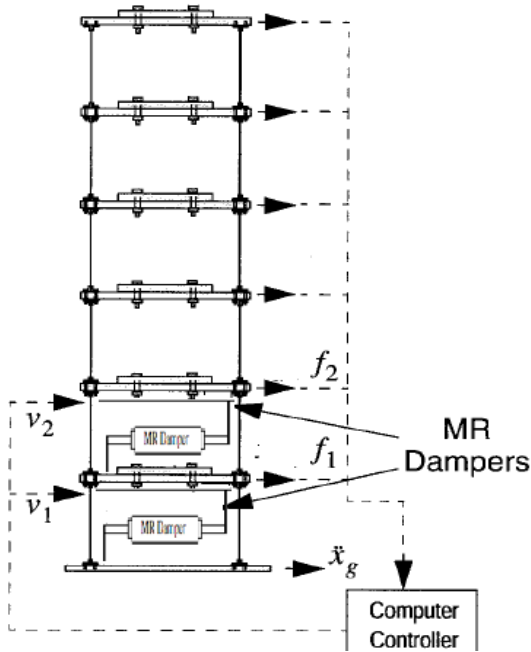


FIG. 4 Schematic of Example

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = -ME\ddot{x}_g + Lu(t) \quad (13)$$

Where,

$M$  = Mass matrix

$C$  = Damping matrix

$K$  = Stiffness matrix

$x(t)$  = Displacement vector of building at time  $t$

$L$  = Location vector for control forces

$u(t)$  = Control force vector at time  $t$

$E$  = Vector of ones

$\ddot{x}_g(t)$  = Ground accelerations at time  $t$

In state space form,

$$A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix}_{2n \times 2n} \quad B = \begin{bmatrix} 0 \\ -M^{-1}L \end{bmatrix}_{2n \times m}$$

$$C = \begin{bmatrix} 0 & I \\ I & 0 \\ -M^{-1}K & -M^{-1}C \end{bmatrix}_{p \times 2n} \quad D = \begin{bmatrix} 0 \\ 0 \\ -M^{-1}L \end{bmatrix}_{p \times m}$$

$$H = \begin{bmatrix} 0 \\ E \end{bmatrix}_{2n \times 1}$$

In simulation the building is subjected to three earthquake history. El-Centro ground motion of 1940 with PGA = 0.349g, Kobe 1995 ground motion with PGA = 0.82g and Northridge 1994 ground motion with PGA = 0.83g. The simulation for semi active control was evaluated for various evaluation criteria as follows:

1. The first evaluation criterion is a measure of the normalized maximum floor displacement relative to the ground, given as  $J_1 = \left( \frac{\text{Max}|x_i(t)|}{x_{\text{unctrl}}} \right)$ , Where  $x_i(t)$  s the relative displacement of the  $i$ th floor over the entire response, and  $x_{\text{unctrl}}$  denotes the uncontrolled maximum displacement.
2. The second evaluation criterion is a measure of the reduction in the interstory drift. The maximum of the normalized interstory drift is,

$$J_2 = \left( \frac{\text{Max}|d_i(t)/h_i|}{d_{\text{unctrl}}} \right), \text{ Where } h_i \text{ the height of each is floor (30 cm), } d_i(t) \text{ is the interstory drift of the above ground floors over the response history, and } d_{\text{unctrl}} \text{ denotes the normalized peak interstory drift in the uncontrolled response.}$$

3. The third evaluation criterion is a measure of the normalized peak floor accelerations, given by  $J_3 = \left( \frac{\text{Max}|\ddot{x}_i(t)|}{\ddot{x}_{\text{unctrl}}} \right)$ , Where the absolute accelerations of the  $i$ th floor,  $\ddot{x}_i(t)$  are normalized by the peak uncontrolled floor acceleration  $\ddot{x}_{\text{unctrl}}$ .
4. The final evaluation criteria considered in this study is a measure of the maximum control force per device, normalized by the weight of the structure, given by



$J_4 = \left( \frac{\text{Max}|f_i(t)|}{f_{\text{unctrl}}} \right)$ , Where is the total weight of the structure (1335 N).

Table 1 shows the normalized responses of different control strategy for the El-Centro earthquake ground motion. The percentage in the parenthesis indicates the response reduction over best passive case. Negative sign indicate the reduction and positive sign indicate increased response as compared to passive on case.

Table 2 shows the evaluated performance indices J1, J2, J3, and J4 for El-Centro, Kobe, and Northridge earthquakes. Results are calculated for passive off, passive on and clipped optimal control (LQR) control strategies.

Table 1 Normalized Controlled Maximum Responses due to Scaled El Centro 1940 Earthquake.

Control Strategy	J1	J2	J3	J4
Passive off	0.863	0.8129	0.902	0.0029
Passive on	0.5039	0.7186	1.039	0.0178
Lyapunov	0.686 (36%)	0.788 (10%)	0.756 (-16%)	0.0178
Decentralized bang bang	0.449 (-11%)	0.791 (10%)	1.00 (10%)	0.0178
Maximum energy dissipation	0.548 (9%)	0.620 (-14%)	1.06 (17%)	0.0121
Modulated homogenous friction	0.421 (-16%)	0.559 (-22%)	1.06 (2%)	0.0178
Clipped optimal + LQG	0.405 (-20%)	0.547 (-24%)	1.25 (38%)	0.0178
Clipped optimal + LQR	0.3289 (-35%)	0.456 (-36%)	0.8809 (-2%)	0.0178

Table 2 Evaluated Performance Indices Due to Various Earthquakes.

Control Strategy	J1	J2	J3	J4
<b>El-Centro</b>				
Passive off	0.863	0.7998	0.9017	0.0029
Passive on	0.5039	0.7071	1.0382	0.0178
Clipped optimal + LQR	0.3289 (-35%)	0.456 (-36%)	0.8809 (-2%)	0.0178
<b>Kobe</b>				
Passive off	0.7514	0.7408	0.847	0.003
Passive on	0.2725	0.2687	0.3757	0.0179
Clipped optimal + LQR	0.3148 (15%)	0.2931 (9%)	0.3842 (2%)	0.0179

Northridge

Passive off	0.8626	0.8747	0.9772	0.0029
Passive on	0.5081	0.6722	0.9635	0.0178
Clipped optimal + LQR	0.5612 (10%)	0.6522 (-30%)	0.8642 (-10%)	0.0178

To compare the performance of the semi-active system to that of comparable passive systems, two cases are considered in which the MR dampers are used in a passive mode by maintaining a constant voltage to the devices. The results of passive-off (0 V) and passive-on (5 V) configurations are included. For the Clipped optimal controller there is no standard method of selecting the Q matrix; therefore, several 12 x 12 Q matrices were arbitrarily chosen and tested. The challenge in clipped optimal controller design is in the selection of Q. Here Q = 8400000 is selected as of benchmark problem.

The peak responses quantities namely, floor displacement, interstory drift and absolute accelerations for each earthquake are tabulated and graphically represented as under.

Table 3 Peak response quantities of the example building under scaled El-Centro 1940 earthquake

Control Strategy	Floor	1	2	3	4	5	6
Uncontrolled	Disp.	0.299	0.572	0.816	1.030	1.205	1.314
	Drift	0.299	0.275	0.255	0.228	0.190	0.112
	Abs acc.	52.059	79.983	92.572	92.137	103.132	147.112
Passive off	Disp.	0.239	0.450	0.673	0.869	1.032	1.133
	Drift	0.239	0.227	0.230	0.211	0.172	0.101
	Abs acc.	62.008	108.142	91.616	97.463	111.615	132.554
Passive on	Disp.	0.130	0.235	0.444	0.573	0.631	0.662
	Drift	0.130	0.106	0.212	0.177	0.149	0.116
	Abs acc.	119.798	130.208	105.547	99.325	91.804	152.615
Clipped Optimal + LQR	Disp.	0.094	0.190	0.259	0.342	0.413	0.432
	Drift	0.094	0.109	0.136	0.123	0.094	0.069
	Abs acc.	129.489	122.465	63.768	60.518	72.832	90.973

Table 4 Peak response quantities of the example building under scaled Kobe 1995 earthquake

Control Strategy	Floor	1	2	3	4	5	6
Uncontrolled	Disp.	1.471	2.875	4.131	5.169	5.915	6.305
	Drift	1.471	1.404	1.261	1.049	0.765	0.406
	Abs acc.	140.388	243.968	323.054	390.174	471.650	530.397
Passive off	Disp.	1.080	2.076	3.005	3.812	4.413	4.738
	Drift	1.080	1.017	0.965	0.832	0.618	0.343
	Abs acc.	174.316	279.199	267.424	318.921	414.225	448.892
Passive on	Disp.	0.342	0.642	1.006	1.343	1.589	1.718
	Drift	0.342	0.301	0.392	0.337	0.265	0.152
	Abs acc.	178.058	192.119	157.603	166.649	171.795	199.100
Clipped Optimal + LQR	Disp.	0.422	0.843	1.234	1.557	1.834	1.985
	Drift	0.422	0.427	0.401	0.354	0.304	0.156
	Abs acc.	169.439	166.229	163.625	196.447	201.402	203.636

Table 5 Peak response quantities of the example building under scaled Northridge 1994 earthquake

Control Strategy	Floor	1	2	3	4	5	6
Uncontrolled	Disp.	0.499	1.024	1.531	1.961	2.267	2.421
	Drift	0.499	0.528	0.508	0.431	0.306	0.155
	Abs acc.	80.370	104.601	108.037	162.948	198.892	202.442
Passive off	Disp.	0.418	0.853	1.283	1.665	1.946	2.087
	Drift	0.418	0.436	0.436	0.385	0.285	0.149
	Abs acc.	109.946	122.131	105.630	134.692	197.393	195.771
Passive on	Disp.	0.257	0.489	0.824	1.062	1.184	1.230
	Drift	0.257	0.235	0.335	0.269	0.237	0.149
	Abs acc.	182.522	139.851	157.731	154.487	163.934	194.636
Clipped Optimal + LQR	Disp.	0.250	0.533	0.807	1.061	1.239	1.358
	Drift	0.250	0.291	0.325	0.256	0.207	0.134
	Abs acc.	156.869	123.363	122.782	142.442	131.664	174.576

The Graphical representation of the responses shown in tables 3, 4 and 5 are represented in figures 5, 6 and 7 respectively.

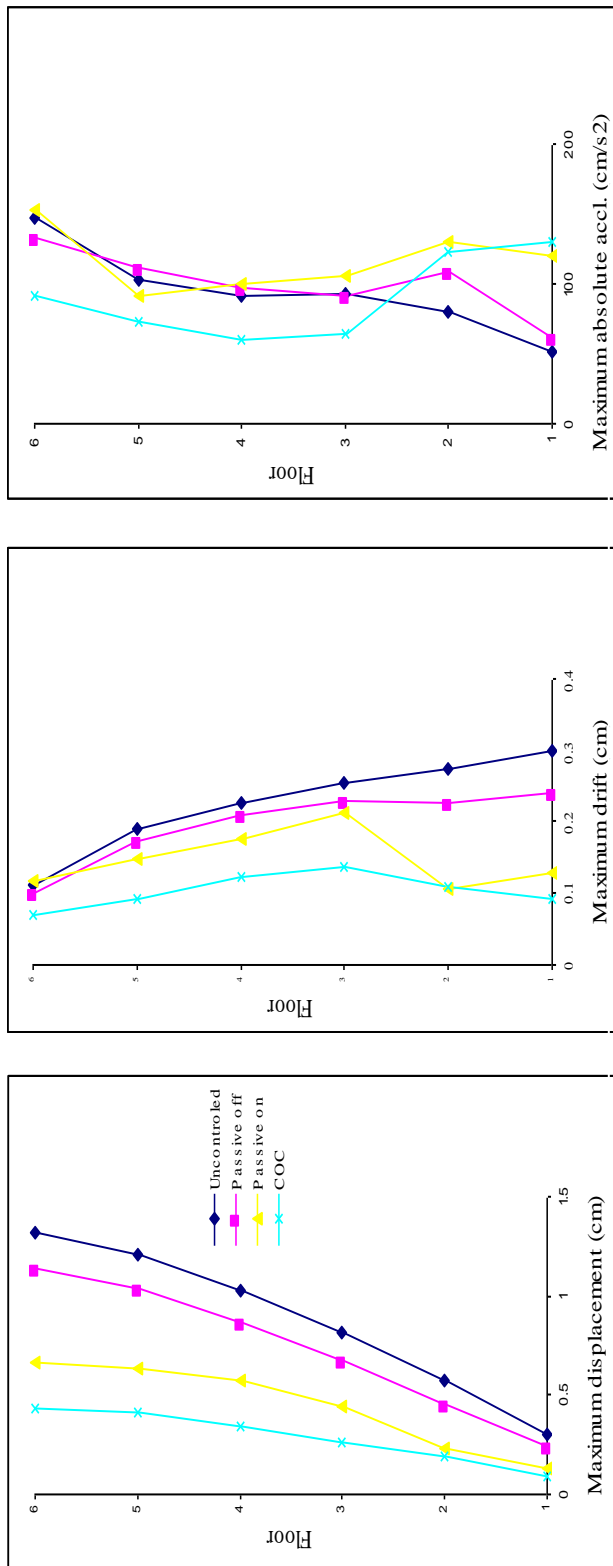


Fig 5 Peak Responses of Each Floor of the Building to the Scaled El-Centro 1940 Earthquake

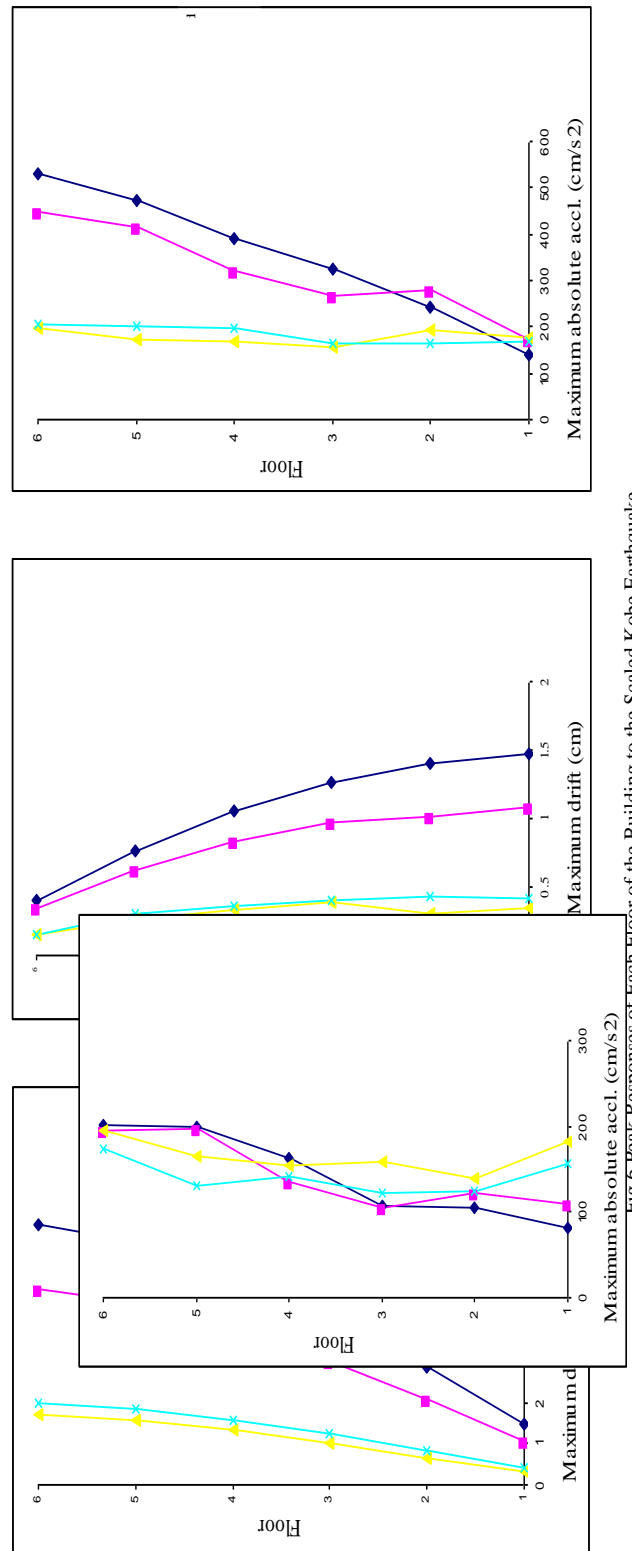
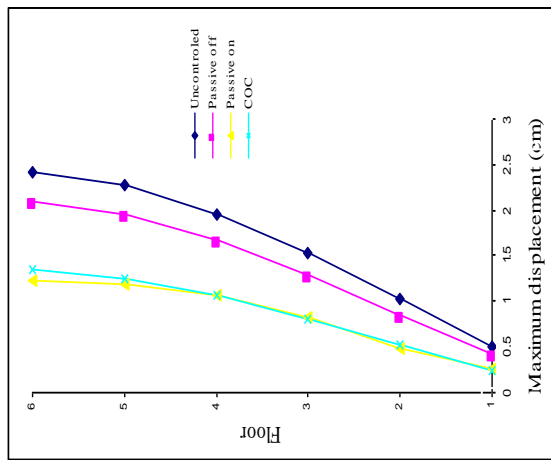


Fig 6 Peak Responses of Each Floor of the Building to the Scaled Kobe Earthquake



Controlled and uncontrolled responses are compared and shown in following figures 8 to 16.

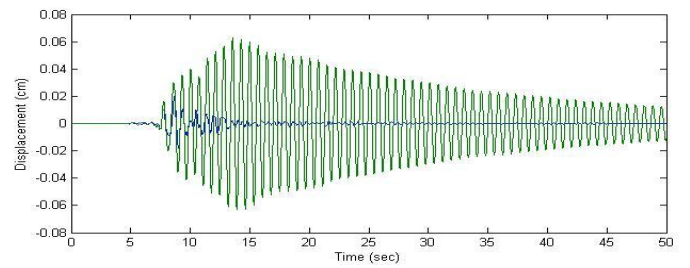


Fig 11 Displacement History under Kobe Earthquake

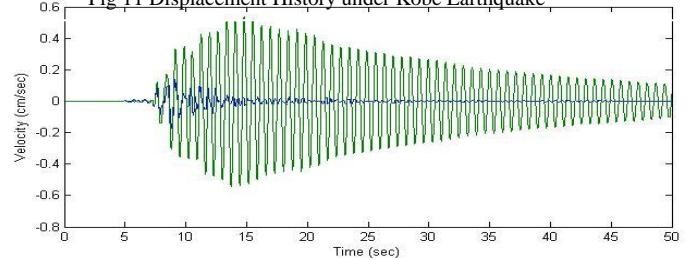


Fig 12 Velocity History under Kobe Earthquake

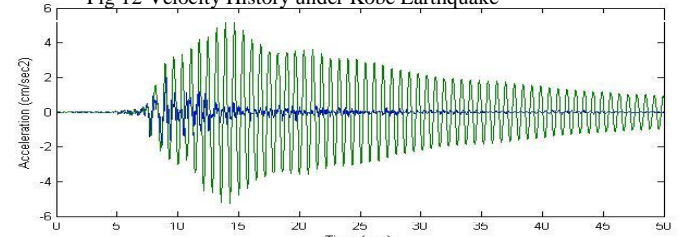


Fig 13 Acceleration History under Kobe Earthquake

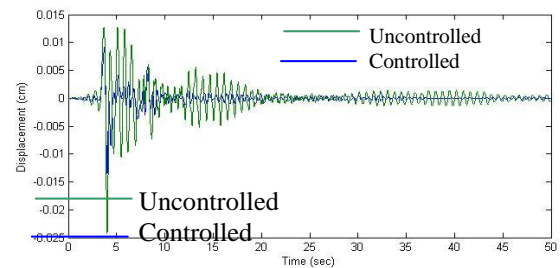
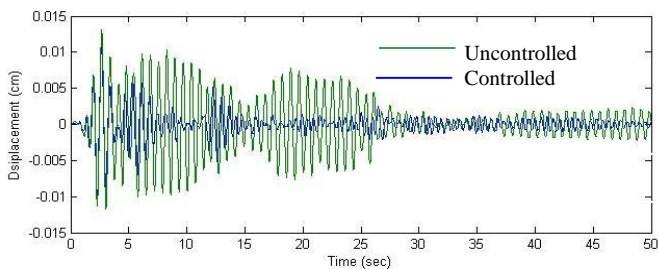


Fig 14 Displacement History under Northridge Earthquake

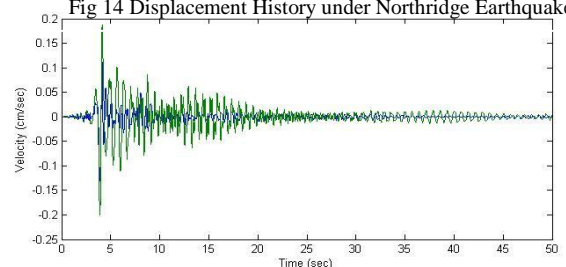
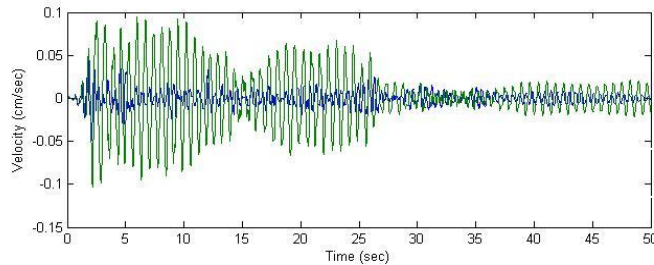


Fig 15 Velocity History under Northridge Earthquake

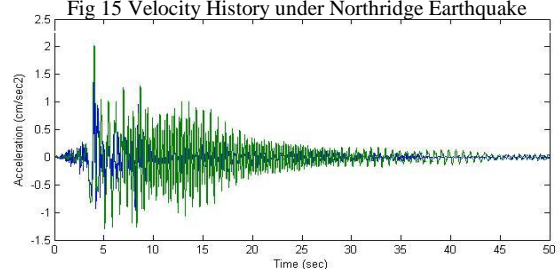
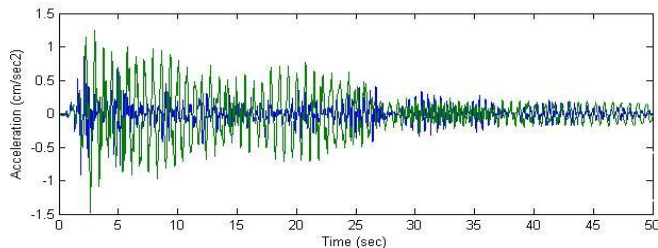


Fig 4.16 Acceleration History under Northridge Earthquake

## VIII. SUMMARY AND CONCLUSION

Semi-active response control of a six-story building equipped with magnetorheological dampers on the lower two floors is studied. Clipped-optimal control algorithm based on linear quadratic regulator (LQR) is employed for semi-active controller design. Bouc-Wen phenomenological model is utilized to investigate the nonlinear behavior of the MR dampers. The performance of the resulting control system is compared to the uncontrolled building through simulation, and the efficiency of using MR dampers with LQR is evaluated for the selected three earthquake ground motions. The results of the study are summarized as follows.

1. The passive-off system reduces the peak floor displacement, peak drift and peak absolute acceleration over the uncontrolled case.
2. The passive-on system further reduces the peak floor displacement and peak drift. However, the peak absolute acceleration increases compared to the passive off system. This is due to the fact that the passive-on system attempts to lock up the first two floors, increases the drift of the upper floors and absolute acceleration of the lower floors of the building.
3. The reduction in normalized maximum floor displacement,  $J_1$ , normalized peak interstory drift,  $J_2$ , and normalized peak absolute acceleration,  $J_3$  are 3%, 19% and 3%, respectively over best passive case is observed using clipped optimal control with LQR.
4. The reduction in peak displacement, peak drift and peak absolute acceleration are **60%**, **63%** and **38%**, respectively over best passive case is found using clipped optimal control with LQR

## REFERENCES

- [1] Chopra, A. K. (2005). "Dynamics of Structures: Theory and Application to Earthquake Engineering," New Delhi: Pearson Education.
- [2] Dyke, S.J., Spencer, Jansen L.M.(1997). "Semi active control strategies for MR damper – A comparative study," Journal of Engineering, Mechanics, Vol. 126, No. 8, pp. 795–803.
- [3] Dyke, S.J., Spencer, B.F., Sain, M, K., Carlson, J.D. (1997). "Phenomenological model of a magnetorheological damper," Journal of Engineering, Mechanics; Vol. 123(3):230\_8.
- [4] Dyke, S.J., Spencer, B.F. (1996). "Seismic response control using multiple MR dampers," In: 2nd international workshop on structural control. Hong Kong University of Science and Technology Research Center.
- [5] Dyke, S.J., Spencer Jr., B.F., Sain, M.K. and Carlson, J.D. (1996). "Modeling and Control of Magnetorheological Dampers for Seismic Response Reduction," Smart Materials and Structures, Vol. 5, pp. 565–575.
- [6] Dyke, S.J., Spencer Jr., B.F., Sain, M.K. and Carlson, J.D. (1997). "A Comparison of Semi-Active Control Strategies for the MR Damper," Proc. of the IASTED Intl. Conf. on Intelligent Info. Systems, pp. 580–584, Bahamas, December 8–10.
- [7] Dyke, S.J., Spencer, B.F., Sain, M, K., Carlson, J.D. (1997). "Phenomenological Model for Magnetorheological Dampers," Journal of Engineering Mechanics, ASCE, Vol. 123, No. 3, pp 230-238.
- [8] Dyke, S.J., Spencer Jr., B.F., Sain, M.K. and Carlson, J.D. (1997). "An Experimental study of Magneto-rheological Dampers for

- Seismic Hazard Mitigation," Proceedings of Structures Congress XV, Portland, Oregon, U.S.A., pp. 58-62.
- [9] Faruque Ali, Ananth Ramaswamy1. (2009). "Hybrid structural control using magnetorheological dampers for base isolated structures," Smart Mater. Struct. 055011 (16pp).
  - [10] Kori, J.G., Jangid, R.S. (2008). "Semi active control of seismically isolated bridges," International Journal of Structural Stability and Dynamics Vol. 8, No. 4 547–568.
  - [11] Maryam, B., Osman, E.O. (2010). "Application of semi-active control strategies for seismic protection of buildings with MR dampers," Engineering Structures 32 (2010) 3040\_3047.
  - [12] MATLAB, (1994). The Math Works, Inc. Natick, Mass.
  - [13] MR Dampers 2006 LORD Technical Data: RD-1005 Damper.
  - [14] Nagrath & Gopal, "Control System Engineering," New age international publication.
  - [15] Spencer, B.F., Nagarajaiah, S. "State of the Art of Structural Control,"
-

# LOW IMPACT DEVELOPMENT MODELLING TO RESTORE URBAN WATERSHED HYDROLOGY IN RAJKOT CITY

A. Ms. VIDHI H. KHOKHANI

**Abstract**—The effects of traditional development practices on the hydrologic cycle have been well documented. Increases in the impervious surfaces associated with urbanization have resulted in increased surface Runoff, decreased baseflow, changes in stream morphology, increased stream temperature, aquatic/riparian loss. The goal of this paper to widely explores concept of low impact development as an innovative approach to stormwater management and ecosystem protection that integrates hydrologic controls into every aspect of site design to mimic the predevelopment hydrologic. This paper presents low impact development modelling of Rajkot city and compare various LID model by using Storm water Management Modelling system 5.0. Finally concluding Best Management Practice for stormwater management.

**Keywords**—Low impact development (LID), SWMM (stormwater management modelling) Bioretention, Grass swales, Predevelopment, Infiltration trenches, Best Management Practices (BMP) .

## I. INTRODUCTION

Low Impact Development is an innovative approach to stormwater management and ecosystem protection that integrates hydrologic controls into every aspect of site design to mimic the predevelopment hydrologic regime. It is not a growth management strategy nor does it rely heavily on density restrictions, rezoning, clustering or conservation measures. Instead, LID focuses on engineering the built environment to maintain ecosystem and hydrologic functions. LID uses new site planning and design principles and a wide array of micro-scale management practices. It is a powerful technology that allows development in a manner that preserves water related ecological functions and maintains development potential. The goal of LID is not to mitigate development impacts, but to recreate and preserve a watershed's hydrologic cycle.

The Stormwater management model EPA SWMM 5.0 is a “dynamic rainfall-runoff simulation model used for single event or long-term (continuous)

simulation of runoff quantity and quality from primarily urban areas” (USEPA, 2009). The model is divided into four conceptual compartments atmosphere, land surface, groundwater, and transportation.

Model is prepared and simulated for the single event rainstorm than it is calibrated and validated for the observed data. Then various model scenario is developed. LID stormwater management practices are capable of reducing the volume of urban runoff and thus have the potential to greatly reduce the size and cost of stormwater basins and conveyance systems (EPA, 2007). Additionally where conventional stormwater BMP's struggle to address water quality issues, LID BMPs are robust. The filtration and infiltration of stormwater provided by many LID BMP's have been shown to reduce dissolved pollutants, nutrients (Davis, Hunt, Traver, & Clar, 2009), and bacteria (Clary, Jones, & Urbonas, 2009). Moreover, the total volume reduction of stormwater - through the process of infiltration - is crucial to maintaining the natural flow regime of urban streams upon which the health of aquatic species depends.

## II. STUDY AREA

The study area is located almost entirely within the boundary of the Rajkot Watershed, and includes forested land, an industrial area surrounding and residential or urbanized areas on the Draft Town Planning Scheme no 19. Fig 1 shows location of Rajkot city and fig-2 shows area selected for study. The study area is divided into 9 small subcatchment and the drainage characteristics of catchment is derived by using SRTM (*Shuttle Radar Topography Mission*) data. A digital elevation model (DEM) with a resolution of 90 m SRTM data available on <http://srtm.csi.cgiar.org/> were used. Figure 3 shows subcatchment area of watershed with rain gauge station and outfall point.



Fig 1 Location of Rajkot city

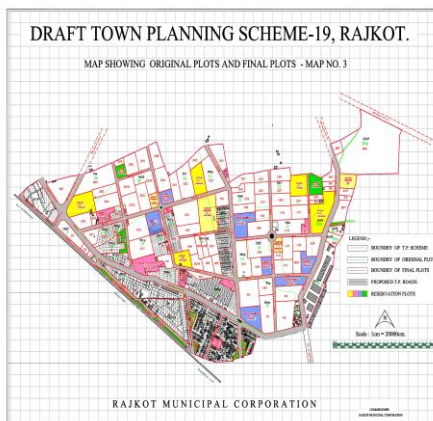


Fig 2 Study Area

The characteristic width and slope were calculated using the methods described in the SWMM Applications Manual (2009) with length and elevation data from the GIS files. Depression storage was taken to be the default value of 0.05 inches, and soil properties were taken from the Stormwater Management Master Plan Table 1 provides a summary of all subcatchment parameters which is derived from using DEM(90 m resolution) available on SRTM website

Table 1 Catchment parameters

Name of subcatchment	Parameters of catchment					Land use%
	Area (ac)	Width ft	Slope %	Impervious area %	Mannings n	
S1	14.5	5671	3.6	85	0.15	Urban-80 Grass area 10
S2	29.3	4738	3.75	8	0.15	Urban-15 Grass -85
S3	31	5079	5.5	5	0.15	Urban -0 Grass 100
S4	42	7366	3.46	2	0.15	Urban 8 Grass 92
S5	34.8	5480	7.5	92	0.15	Urban 98 Grass 2
S6	40	6688	3.6	25	0.15	Urban-30 Grass70
S7	57	7498	4.75	56	0.15	Urban-50 Grass-50
S8	50	6900	5.9	30	0.15	Urban-40 Grass-60
S9	24.5	554	6.2	92	0.15	Urban 98

Name of subcatchment	Parameters of catchment					Land use%
	Area (ac)	Width ft	Slope %	Impervious area %	Mannings n	
		6				Grass 2

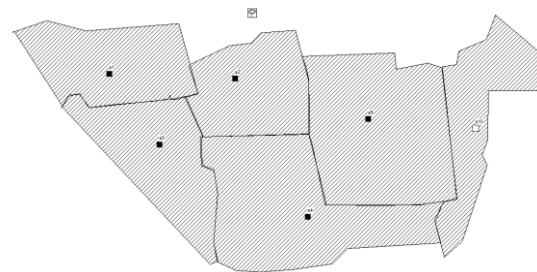


Fig 3 Subcatchment

### III. MODEL SENARIO

Most commonly used LID are described as follows in Table 2

Bioretention/ Rain Gardens
Disconnection of Impervious Areas
Dry Wells
Filter Strips
Grassed Swales/ Bioretention Swales
Infiltration Trenches
Permeable Pavement
Rain Barrels & Cisterns
Reducing Impervious Areas
Soil Amendments
Tree Box Filters
Vegetated Buffers
Vegetated Roofs

Table 2 Common LID BMPs

Impervious surfaces are responsible for more stormwater runoff than any other type of land use. Paved surfaces that often replace vegetated areas increase the volume and frequency of rainfall runoff. In addition, these surfaces provide a place for pollutants to accumulate between rainfall events, and are later washed off into receiving waters. Keeping runoff on-site to allow for infiltration as well as chemical, physical and biological processes to take place is the most effective means of reducing pollutant loadings. This study quantifies how much runoff and pollutant loadings can be reduced by using swales and landscaped depressions in parking lots. In addition to investigating basins with and without swales, were compared. The paper is describe swales , infiltration trenches and rain gardens as solution of excessive runoff.figure shows effect of urbanization over the area

Table 3 Models senario
1. Predevelopment
2. Traditional Curb and Gutter with DCIA:
3. LID 1 (Full Infiltration) (Infiltration trenches,rain barrels,pervious pavement)
4. LID 2 (Partial Infiltration with Underdrains) (Rain water harvesting, Swales,bioretention swale)

IV: MODEL CALLIBRATION AND VALIDATION

The model was calibrated within S5, where field measured stream stage and discharge data was available for comparison. This reach was chosen for calibration because it represents the input to the S1, where quantity and quality conditions are of primary concern to the Stormwater Team. Goals for calibration of the model included the shape and timing of the hydrograph, the total flow volume under the hydrograph, and the peak discharge of the hydrograph. Calibration was performed by visual comparison of field measured and SWMM predicted hydrographs. The first step in the calibration procedure involved determining a rating curve for s1. A rating curve had been created from field gathered stage and discharge data, but it was found to be inconsistent with the rating curve predicted by SWMM. Further, although the field rating curve appeared to fit the field data well with an R2 value of 0.8986, it exhibited substantial percentage errors with respect to field measured discharge for depths above 0.4 feet To eliminate this inconsistency, the SWMM rating curve was used to calculate new discharges corresponding to each field measured stage value. These new discharge values could then be compared to those predicted by SWMM for calibration.

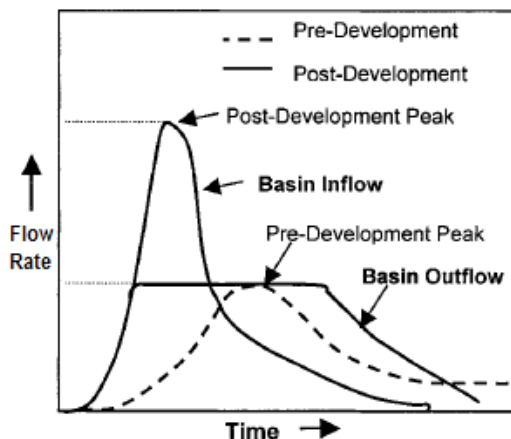


Fig. 4 Effect of urbanization over area

V: MODEL RESULT

Simulating various model following results is obtained for different catchments.

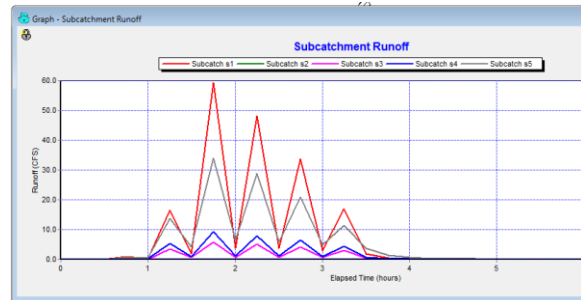


Fig.5 Runoff for catchment 1,2,3,4,5

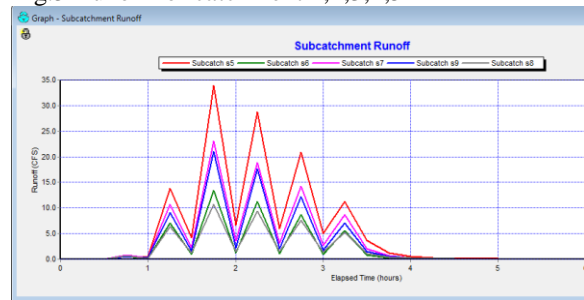


Fig. 6 Runoff for catchment 5,6,7,8,&9

TABLE 4 RUNOFF AFTER LID CONTROL PROVIDED

Table - Subcatch s1		
Days	Hours	Runoff (CFS)
0	00:15:00	0.00
0	00:30:00	0.00
0	00:45:00	0.97
0	01:00:00	0.34
0	01:15:00	16.57
0	01:30:00	1.90
0	01:45:00	59.30
0	02:00:00	3.78
0	02:15:00	48.17
0	02:30:00	3.66
0	02:45:00	33.68
0	03:00:00	3.02
0	03:15:00	17.01
0	03:30:00	1.83



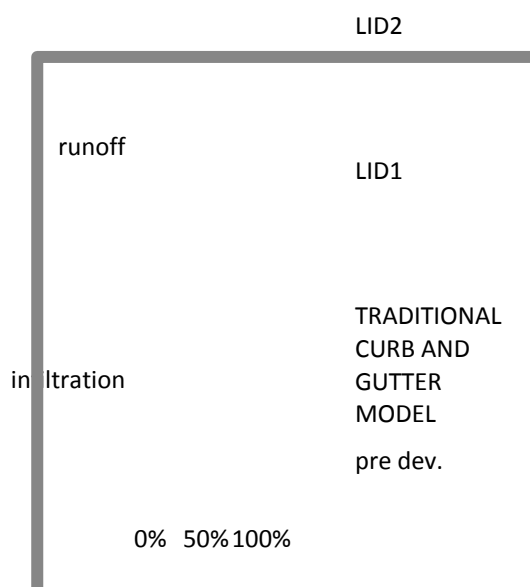
VI: CONCLUSION

Follow charts and figure shows increased amount of infiltration and baseflow. Also shows that the LID 1 & LID 2 model restores runoff and infiltration volume as before urbanization.

Table 5 Design storm Result

Models	Peak Discharge (cfs)	Runoff Volume (cfs)	Infiltration Volume cfs	% Runoff
Pre development model	19.4	1.03	8.97	12%
Traditional curb and Gutter Model	53.0	4.94	3.06	61.7%
LID 1	24.9	1.78	6.24	22.2%
LID 2	24.2	1.65	6.36	20.7%

Fig 6 Infiltration & Runoff Volume



The results demonstrate that LID can be used exclusively to meet stormwater requirements in given area. The water quality requirement was met by routing each impervious area through a LID Best Management Practices which was sufficiently sized to capture the WQCV. Both LID models also demonstrated that they were able to maintain the predevelopment peak flow rates. Normally this would be considered sufficient to meet most municipal stormwater requirements. However, traditional stormwater management does not fully maintain the predevelopment hydrologic regime; it controls peak flow rates while paying little attention to the increased flow duration and volume of runoff

associated with urbanization. To address this, the LID models were built using both peak flow and water balance as criteria. The result was that the LID 1 model replicated the predevelopment flow regime for the 2-, 10-, and 100-year design storms very well and acceptable. The LID 2 model, which includes underdrains, also met the stormwater requirements, but did not replicate the shape of the Rajkot predevelopment hydrographs as closely as the LID 1 model did.

In the historical simulations, both LID models followed the predevelopment peak flow frequency and flow duration curves very well for the lower frequency flows, and split the difference between the DCIA and NDCIA curves for the higher frequency flows. In contrast, the Traditional model with Normal Control and with Over Control could not replicate the predevelopment water balance and significantly prolonged the duration of high frequency flows, revealing that it could not completely replicate the predevelopment hydrologic regime. This is where LID proved to have the upper hand and showed that it is more effective at controlling stormwater than traditional management practices.

These modeling results hinge on one major assumption, that the rate of infiltration through the permeable pavement surface is not a limiting factor. Literature on this subject suggests that when permeable pavements are maintained, the assumption is valid.

An interesting and valuable SWMM modeling finding is that the two different methods used to model LID performed very similarly for short duration storms but different for long duration storms. They performed nearly identically. However, the LID 2 model yielded much lower peak flow rates than both the LID 1 and predevelopment models. In the historical simulations, the LID 2 curve followed the LID 1 curve very closely for all but the very low flows. These findings suggest that for shorter duration storms, the full infiltration method (LID 1) could be used to model partial infiltration BMPs (LID 2). This is valuable because the modeling method used in the LID 2 model is much more complicated and less stable. The result would be that peak flows are slightly over estimated and the runoff volume is slightly under estimated.

The cost comparison evaluated whether it is financially practical to manage stormwater exclusively with LID. This comparison revealed that the LID construction cost is about twice as much as the traditional storm drain system; this comparison included the road construction cost because the permeable pavement converts the road to a functional part of the drainage system. While this would seem to be cost prohibitive, a further investigation found that in urban areas this cost could be recovered by developing the detention basin land, which is no longer needed because of LID. It also revealed that in this site design, over half of the total LID cost was in

the permeable pavement surface itself and thus the profitability of the project would be very sensitive to the construction cost of permeable pavement. Furthermore, the extensive underdrain system, in the LID 2 model, was relatively expensive and recovering this extra cost could be difficult. There are situations in which LID can be cost effective, as well as times in which the cost will outweigh the benefits of LID. If the cost of permeable pavement could be lowered, LID could more rapidly become a viable option for developers.

## VII: CONCLUSION

This paper found that LID practices alone can meet the storm drainage requirements. However rather than using LID exclusively, a more practical option may be using LID to address water quality and to match the predevelopment water balance, while using a detention pond to restore predevelopment peak flows. In three of the four LID 1 models, the peak rate - not water balance - controlled the sizing of the pavement base depth. In fact, the LID 1 Low Infiltration model required less than half of the pavement storage to match the predevelopment water balance in both locations. Once the predevelopment water balance has been met, the size of the detention pond needed to control the 100-year peak flow should be substantially reduced. Quantifying the size of that pond was beyond the scope of this thesis; however it would be about 30-40% of the traditional scenario pond size. This option would significantly reduce the amount and cost of permeable pavement, and would free up a majority of the detention pond area for development.

LID is more complicated to model, design, construct, and maintain than its conventional counterparts. Care should be taken when planning and designing LID systems and those considering the use of LID should first count the cost and decide whether it is a worthwhile investment. While the benefits of LID are immense, it is not a 'silver bullet' for stormwater management.

## REFERENCES

- [1] Atlanta Regional Commission (2001). Georgia Stormwater Management Manual. (Vol. 2: Technical Handbook)
- [2] Baker, D., Pomeroy, C., Annable, W., MacBroom, J., Schwartz, J., & Gracie, J. (2008). Evaluating the Effects of Urbanization on Stream Flow and Channel Stability --- State of Practice.
- [3] California Stormwater Quality Association (CASQA) (2003) Infiltration Trench. California Stormwater BMP Handbook - New Development and Redevelopment: [www.cabmphandbooks.com](http://www.cabmphandbooks.com).
- [4] City of Fort Collins (1997). Storm Drainage Design Criteria and Construction Standards. Fort Collins, CO
- [5] Clary, J., Jones, J., & Urbonas, B. (2009). Challenges in Attaining Recreational Stream Standards for Bacteria: Setting Realistic Expectations for Management Policies and BMPs. Paper presented at the World Environmental and Water Resources Congress 2009: Great Rivers
- [6] Davis, A., Hunt, W., Traver, R., & Clar, M. (2009). Bioretention Technology: Overview of Current Practice and Future Needs. *Journal of Environmental Engineering*, 135(3), 109-117.
- [8] Department of Defense (DOD) (2004) Unified Facilities Criteria - Low Impact Development: U.S. Department of Defense, Washington, D.C.
- [9] EPA (2000). Low Impact Development - A Literature Review. U.S. Environmental Protection Agency, Washington, DC
- [10] EPA (2005). Stormwater Phase II Final Rule. U.S. Environmental Protection Agency, Washington, DC.

# Behavior of Single Flanged Pile Foundation

A. Ankit M. Bhalodi,

B. A. K. Verma

C. Darshana Bhatt

**Abstract**-Piles are used to transmit foundation load through soil strata of low bearing capacity to rock strata of high bearing capacity. Pile with flanges is not generally used in present day but the use of pile with flanges can decrease the deflection and it also reduces the settlement. Flanged pile foundation increase the load carrying capacity of foundation some of the load carried by flange and some of the load carried by pile. The present study attempts to do a two dimensional finite element of Single pile and Flanged pile foundation in single layered sand. Flange along with pile is provided at different location and the comparison of result is done.

**Keywords**-Single pile, Flanged pile, Finite element analysis, Settlement.

## I. INTRODUCTION

One of the most important aspects of a civil engineering project is the foundation system. Designing the foundation system carefully and properly, will surely lead to a safe, efficient and economic project overall. In other words, foundation system design is one of the most critical and important step when a civil engineering project is considered. Until quite recently, there were some separately used systems like shallow foundations such as rafts and deep foundations such as piles[1].

In the past few years, there has been an increasing recognition that the use of piles to reduce settlements that can lead to considerable economy without compromising the safety and performance of the foundation[3].

Piles are relatively long, slender members that transmit foundation load through soil strata of low bearing capacity to deeper soil or rock strata of high bearing capacity. The load transfer may be by vertical distribution of the load along the pile shaft or a direct application of load to a lower stratum through the pile point. A vertical distribution is made using a friction pile and a direct load application is made by a point or end bearing pile. Piles are used when it is desirable to transmit loads to strata beyond the practical reach of shallow foundations for economic, constructional or soil condition considerations.

In addition to supporting structures, piles are also used to anchor structures against uplift forces and to assist structures in resisting lateral and overturning forces[6].

A. Studying at B.V.M Engineering college, M.E. Structural Engineering, Vallabh Vidyanagar, Gujarat, India, (email-apatelb@gmail.com)

B. Professor and head of structural engineering department with B.V.M Engineering college Vallabh Vidyanagar, Gujarat, India, (email-akvbvm@yahoo.co.in)

C. Associate Professor with B.V.M Engineering college, Vallabh Vidyanagar, Gujarat, India, email-darshanabvm@gmail.com)

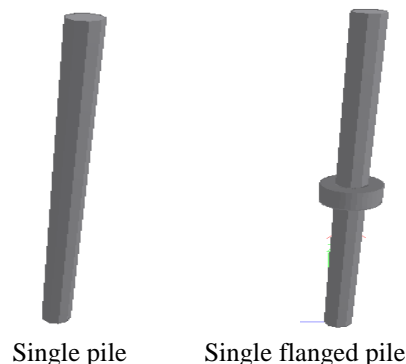
Raft foundations are used to spread the load from a structure over a large area, normally the entire area of the structure. They are used when column loads or other structural loads are so large that more than 50 percent of the area is covered by conventional spread footing and individual pad foundations would nearly touch each other. A raft foundation normally consists of a concrete slab which extends over the entire loaded area. They are often needed on soft or loose soils with low bearing capacity, as they can spread the loads over a larger area[6].

When raft foundation alone cannot provide sufficient stability then combination of piled raft foundation is used. The use of a limited number of strategically located piles with raft can improve the ultimate load capacity and reduce the settlement and differential settlement. It is used particularly for supporting the load of high buildings or towers[3].

As shown that raft is used along with number of pile to reduce the settlement same way flanges along with pile used to reduce the settlement.

To improve the pile foundation the flanges along with pile can also be used. One or more number of flanges can be provided in pile. The use of pile with flanges is very useful in reducing settlement. Flanged pile foundation is a new technique which is not been used in field. By using flange along with pile can increase the load carrying capacity as some of the total load carried by flange also.

The aim of this paper is to describe a finite element analysis of deep foundations: single piled and Flanged pile foundations. A basic parametric study is firstly presented to determine the influence of single layered sand bed on foundation and comparison of result of single pile and Flanges along with pile at different location. The behavior of single piled and Flanged pile foundations is analyzed in more details using axisymmetric models in finite element analysis program(PLAXIS).



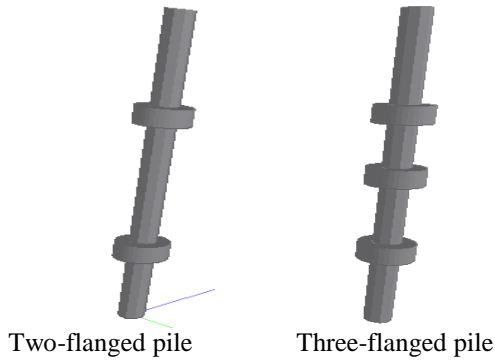


Fig. 1. Single pile and flanged pile foundation

## II. LITERATURE REVIEW

Pile is extensively used to support tall buildings. Several researchers have carried out studies to predict the behavior of pile foundation. Zakia Khelifi, A. Berga and N. Terfaya (2011) presented on behavior of axially and laterally loaded pile foundation. Negative skin friction and ground water table is an important factor of pile foundation[8]. C. J. Lee and C. R. Chen (2003) presented on negative skin friction on piles due to lowering of groundwater table. It was found that negative skin friction force decreased with decreasing spacing to diameter ratio of piles and with increasing number of neighboring piles[2].

In recent times finite element analysis is being extensively used for modeling various foundation systems including piled raft foundation. This requires the discretization of both structural foundation system and soil. Prakoso and Kulhawy (2001) analyzed the behavior of vertically loaded piled raft foundation using elastic and elastic-plastic finite element models which involved the analysis of a three dimensional piled raft as a two dimensional strip piled raft[1]. J. Sebastian (2008) investigated on interface strength factor (Rinter) of single pile and piled raft foundation in PLAXIS. It was found that interface strength factor has a influence on settlement[3]. Sona Vasudev (2009) presented on the behavior of piled raft foundation in homogeneous and layered sand. It was found that there is considerable reduction in settlement of raft by increasing pile length[6]. H. Seo, D. Basu and R. Salgado (2009) presented on load-settlement response of rectangular and circular piles in multilayered soil using finite element program ALPAXL. It was found that in case of pile in multilayered soil, pile response depends on soil layering, with uppermost soil having the most effect on pile head stiffness[4]. H. Chik(2009) presented on lateral behavior of single pile in cohesion less soil subjected to both vertical and horizontal loads using finite element program PLAXIS 3D[5].

## III. FINITE ELEMENT ANALYSIS

The finite element analysis of single pile and flanged pile foundation was undertaken using the finite element program PLAXIS V8. Development of PLAXIS began in 1987 at the Technical University of Delft as an initiative of the Dutch Department of Public Works and Water Management. The initial goal was to develop an easy-to-use 2D finite element code for the analysis of river embankments on the soft soils of the lowlands of Holland. In subsequent years, PLAXIS was extended to cover most other areas of geotechnical engineering. Because of continuously growing activities, a company named PLAXIS b.v. was formed in 1993. In 1998, the first PLAXIS version for Windows was released. In the mean time a calculation kernel for 3D calculations was being developed. After several years of development the PLAXIS 3D Tunnel program was released in 2001.

The two dimensional finite element program was used in this work. In this program, modeling was carried out in axisymmetric conditions. Axisymmetric condition is used for circular element and plain strain condition is used for square and rectangle element. The soil was modeled by triangular element with 15 nodes and elastoplastic law behavior obeying the Mohr-Coulomb failure criterion. The pile was modeled as an elastic linear material. 15- nodes triangle element is very accurate element that has produced high quality stress results.

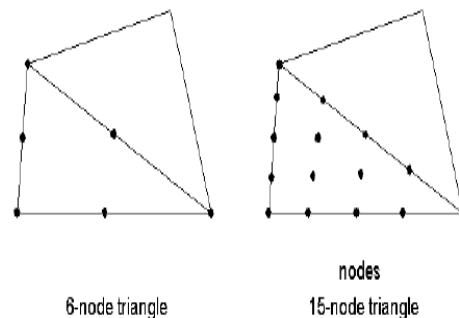


Fig. 2. Triangular element of finite element analysis program

### A. Single Pile Modeling

A working area of 25 m width and 40 m depth has been used. At the axis of symmetry the pile has been modeled with a length of 18 m and a diameter of 1 m. The soil bed is modeled as a single layer of sand. The properties of sand and concrete are described in table-1. The ground water is located at 40 m below the soil surface. In this way we did not take into account the water influence. Along the length of the pile an interface has been modeled. We extended this interface to 0.5 m below the pile inside the soil body to prevent stress oscillation in this stiff corner area. A typical application of interfaces would be to model the interaction between a sheet pile wall and the soil, which is intermediate between smooth and fully rough. The roughness of the interaction is modeled by choosing a suitable value for the strength reduction factor

in the interface ( $R_{inter}$ ). Load applied on pile foundation was prescribed displacement load and prescribed displacement was given as 0.1m. That mean finite element program will apply the load until it reaches the settlement of 0.1m.

Material used in analysis was dense sand hence it was selected under drained condition.

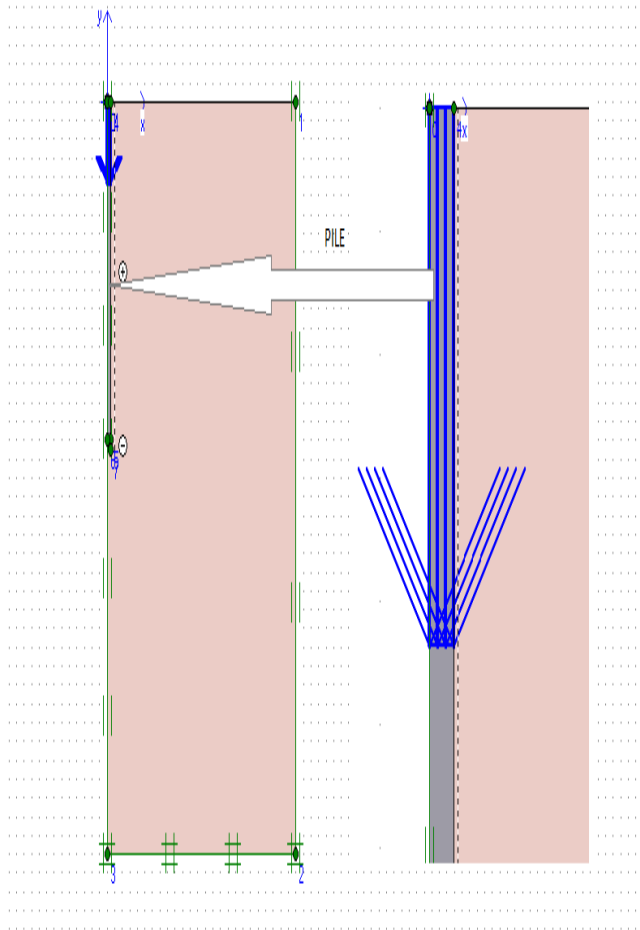


Fig. 3. Global geometry of the axisymmetric model of the single pile

TABLE 1  
MATERIAL PROPERTIES

Parameter	Symbol	Sand	Concrete	Unit
Material Model	Model	Mohr-Coulomb	Linear Elastic	-
Unsaturated weight	$\gamma_{unsat}$	19	25	$kN/m^3$
Saturated weight	$\gamma_{sat}$	21	25	$kN/m^3$
Permeability	k	1	0	m/day
Young's modulus	$E_{ref}$	50000	$30000 \times 10^3$	$KN/m^2$
Poisson's ratio	$\nu$	0.3	0.2	-
Cohesion	$C_{ref}$	1	-	$KN/m^2$
Friction angle	$\Phi$	38	-	$^\circ$
Dilatancy angle	$\psi$	0	-	$^\circ$
Interface factor	$R_{inter}$	0.67	Rigid	-

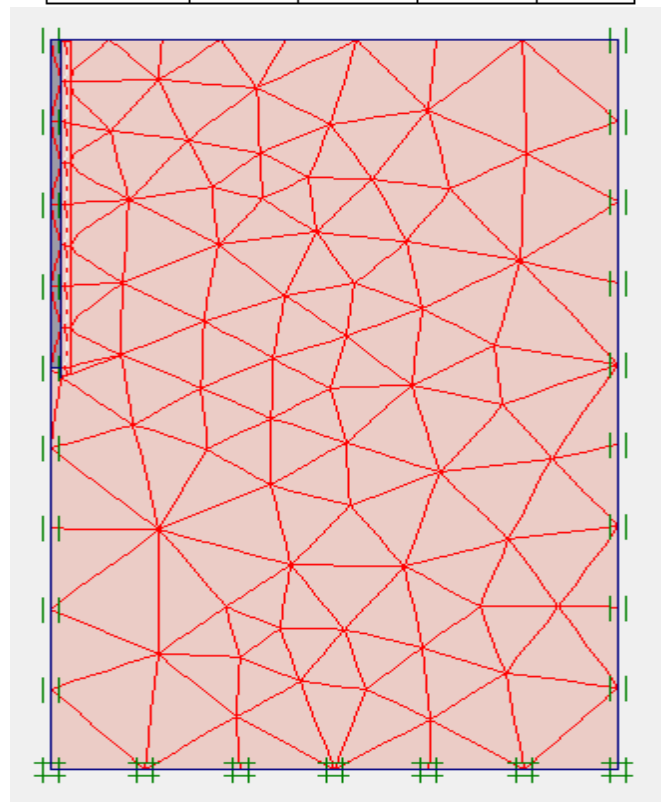


Fig. 4. Mesh Generation of pile foundation

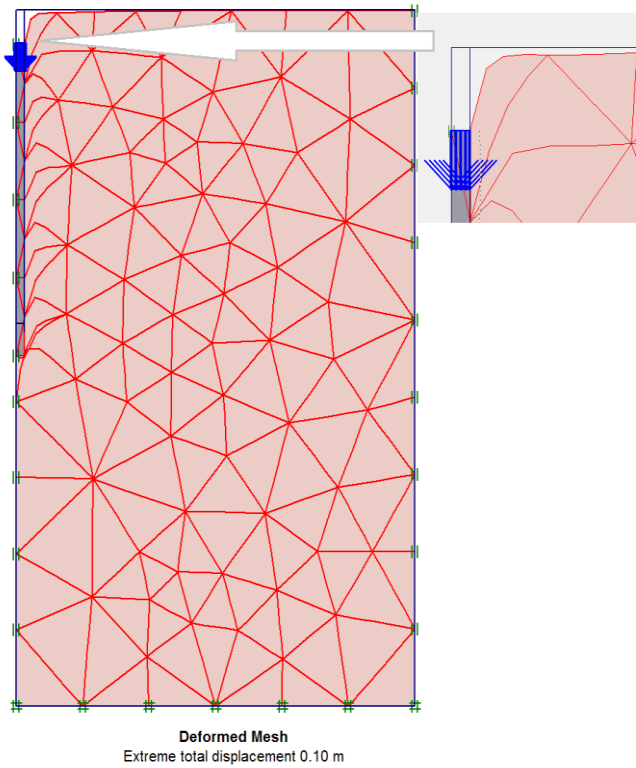


Fig. 5. Deformed Mesh of pile foundation

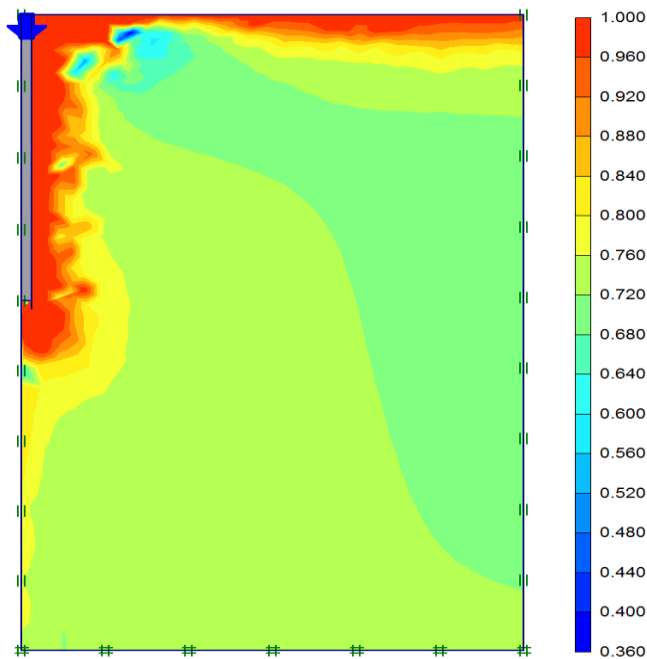


Fig.6. Relative Shear Stresses of single pile foundation

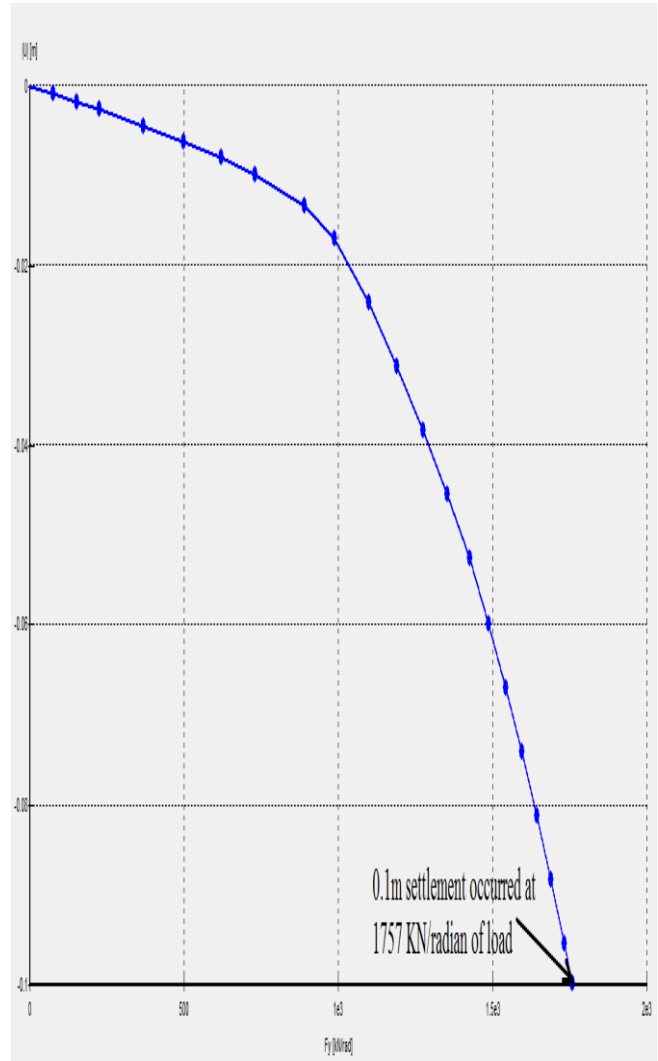


Fig. 7. Settlement curve of single pile

**B. Flanged Pile Modeling**

A working area of 25 m width and 40 m depth has been used. At the axis of symmetry the pile has been modeled with a length of 18 m and a diameter of 1 m. Flanges has been modeled with 1 m depth and 2 m diameter. The number of flanges has been used is one at different location which is 0.1L(L is length of pile), 0.2L, 0.3L, 0.4L, 0.5L, 0.6L, 0.7L, 0.8L and 0.9L. The soil bed is modeled as a single layer of sand. The properties of sand and concrete are described in table-1. The ground water is located at 40 m below the soil surface. In this way we did not take into account the water influence. Along the length of the pile an interface has been modeled. We extended this interface to 1 m below the pile inside the soil body to prevent stress oscillation in this stiff corner area.

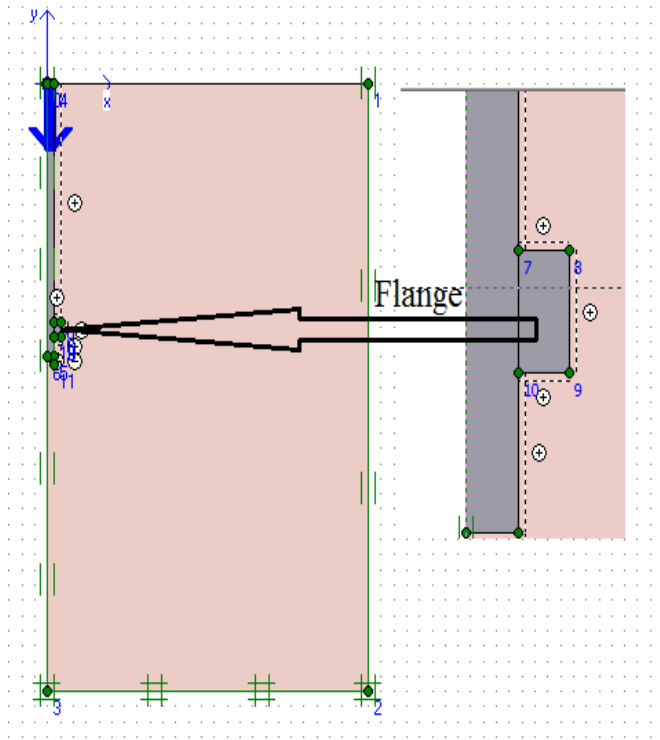


Fig. 8. Modeling of 0.9L flanged pile foundation

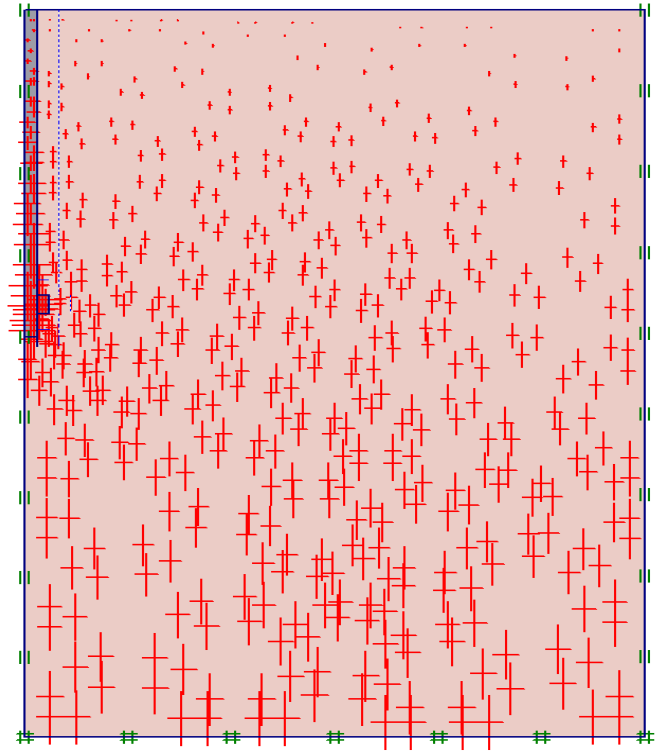


Fig.10. Effective Stresses(Extreme Effective stress is-862.73 KN/m<sup>2</sup>)

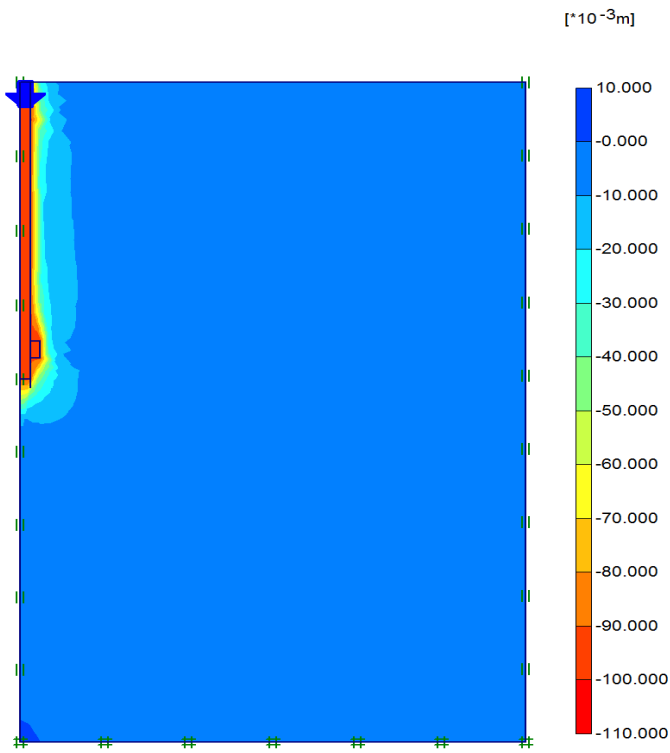


Fig.9. Vertical Displacement of one flanged pile with flange at 0.9L

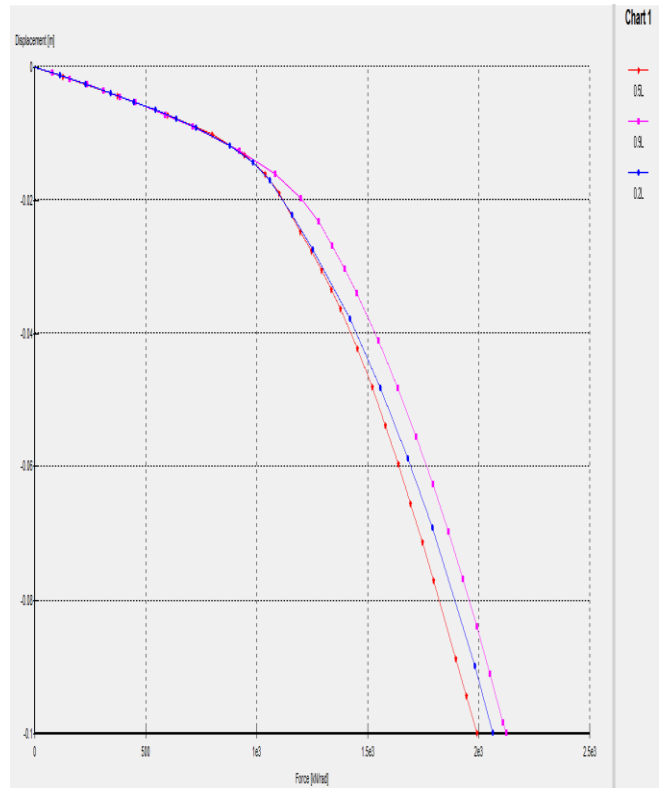


Fig. 10. Load-settlement curve of flanged pile foundation where location of flange are 0.2L, 0.5L and 0.9L

Table 2  
Result of Flanged Pile Foundation with Different Location of Flange

Position of Flange	Load in KN/radian at which 0.1m Settlement Occurred
0.1L	2024
0.2L	2064
0.3L	2031
0.4L	1988
0.5L	1993
0.6L	1867
0.7L	1931
0.8L	1941
0.9L	2125

CONCLUSION

- With the use of flange along with single pile it is found that it reduces the settlement of pile. It is also found that by increasing number of flanges there is a more reduction in settlement.
- Flanged pile foundation with one flange increases the load carrying capacity by 20% than single pile foundation.
- Flanged pile foundation with one flange reduces the settlement by 40% than single pile foundation.
- There is not much difference in result by providing flange at top and bottom of flange.
- There is not such benefit in providing flange at middle of pile foundation.

REFERENCES

[1]W. Prakoso, F. Kulhawy, “Contribution to piled raft foundation design” Journal of geotechnical and geoenvironmental engineering, vol-127, 2001

[2]C. Lee and C. Chen “Negative skin friction on piles due to lowering of groundwater table”, Journal of the southeast asian geotechnical society, April 2003

[3]J. Sebastien, “FE-Analysis of piled and piled raft foundation”, Conference on Soil Mechanics and Foundation Engineering at Graz University of Technology, August-2008

[4]R. Salgado, H. Seo and D. Basu “Load-Settlement response of rectangular and circular piles in multilayered soil”, Journal of Geotechnical and Geo-environmental Engineering, Vol.135, No. 3, March 2009

[5]J. Abbas, Z. Chik and Q. Mohammed “Lateral behavior of single pile in cohesionless soil subjected to both vertical and horizontal loads”, European Journal of Scientific Research Vol.29 No.2, 2009

[6]S. Vasudev “Behavior of piled raft foundation in homogeneous and layered sand”, 10th National Conference on Technological Trends, 6-7 Nov 2009

[7]B. Yilmaz “An analytical and experimental study on piled raft foundations”, journal of graduate school of national and applied sciences of middle east technical university, February 2010

[8]Z. Khelifi, A. Berga and N. Terfaya “Modeling the behavior of axially and laterally loaded pile with a contact model”, Electronic journal of geotechnical engineering Vol.16, 2011

[9]J. Horvath “Static analysis of axial capacity of driven piles in coarse-grain soil” Conference of Manhattan College School of Engineering Center for Geotechnology, June 2002

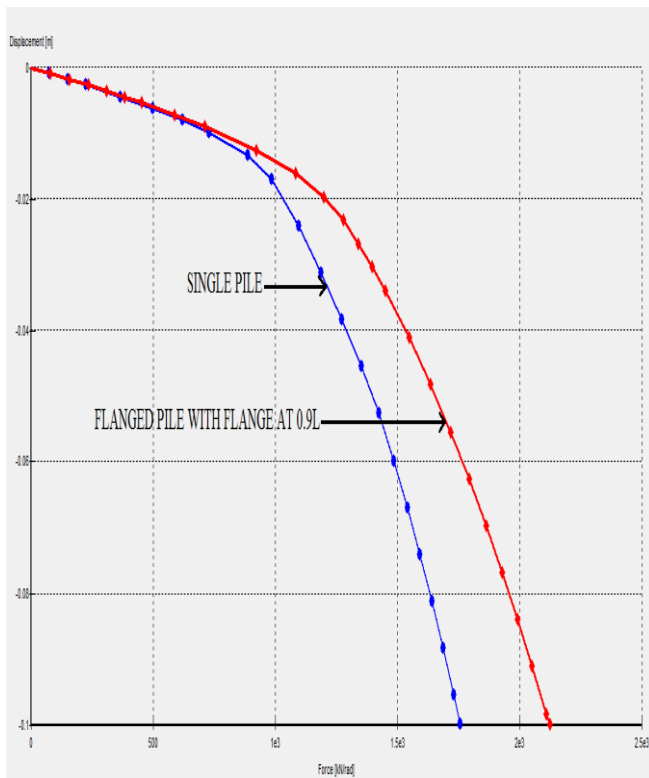


Fig. 11. Load-Settlement curve of single pile and flanged pile foundation



# STUDY OF PROCESS PARAMETERS ON ELECTROKINETIC REMEDIATION

A. Prof. Ashok C.Patel

B. Mr. Kandarp.K.Thaker

C. Prof.Vrundani Vaidya

**Abstract**— Electro kinetic remediation is the method for removing heavy metal Contaminants from fine grained soils especially with low permeability, where other methods do not remain effective. In this process DC is passed through an array of embedded electrodes in the soil. Charged ions are moved toward the electrodes due to their polarity. Uncharged contaminants such as soluble organic molecules move with the bulk liquid by process of electro osmosis. Contaminants reaching the electrode reservoirs can be easily pumped out and treated.

**Keywords**-Clays, electrokinetic remediation, heavy metal ,electrode type, electrode spacing, Chromium

## I. INTRODUCTION

**D**ue to continuous chemical industrial growth in south Gujarat and Saurashtra region heavy metals are observed. These industrial areas contain industrial units including dye factories, textile, rubber, pesticide and paint manufacturers, pulp and paper producers, pharmaceutical, engineering and chemical companies. The Ankleshwar Industrial Association has estimated that its members generate between 250-270 million liters a day of liquid waste, and approximately 50 thousand tones of solid waste annually (Bruno 1995). Of this, 58% arises from the manufacture of dyes and dyes intermediates, 19% from drugs and pharmaceuticals, and 5% from inorganic chemicals (CPCB 1996). Dedicated freight corridor will augmented Industrialization & there by pollution. Recently Government has announced development of polluting industry in Dahej and Dholera.

Heavy Metals contaminants like mercury, cadmium, lead, nickel, copper, zinc, chromium, and manganese may be generated in these regions. In all these region soil observed is fine Grained fat clay and therefore contamination if any, shall be difficult to remove by using conventional methods. In those circumstances, Electro kinetic remediation technique is becomes important and hence a need is felt to study the same systematically.

A.Asst. Prof., Civil Engg. Dept. HGCE, Vahelal, A'bad  
(patelashok58@yahoo.com)

B.Director, K.C.T Consultancy Services, Plot no.1, Sayona Silver Estate-  
Part II, Gota, A'bad- 382481 9kctconser@yahoo.com)

C. Asst. Prof., Civil Engg. Dept. HGCE, Vahelal, A'bad  
(vrundanivaidya@gmail.com)

## II. METHODS OF SOIL REMEDIATION

### A. Soil Washing (Physical-Chemical):

This method is fast and highly effective for cleaning up strongly contaminated soils. This method provides practically total removal of contamination from soil. However, high costs of cleaning installation and utilization are associated with that. Generation of a large amount of solid and liquid wastes that need later management is again a problem in it self. Necessity for transportation of the clean-up soil requires lot of care and additional cost

### B. SOIL FLUSHING

This in situ method is relatively low invasive and without soil removal. Solid wastes is not generated and therefore community acceptance is higher than in the case of soil washing but, a large amount of liquid and semi-liquid wastes are generated. Incomplete removal of contaminants especially strongly bounded heavy metals remains in the soil.

### C. Chemical methods:

The method has broad spectrum of applicability, High efficiency, High specificity of application for individual pollutants, nevertheless associated costs of application is very high. Additionally, large amount of wastes is produced, including hazardous waste. There are problems of process control, especially in the case when used in-situ.

### D. Electro kinetic Remediation

Electro kinetic remediation is an innovative method for removing especially heavy metal contaminants from soil. It is a developing technique which is used for an extraction of heavy metals from the contaminated soil. The electrokinetic method is most effective when cleaning polluted sand and sandy loam.

## III. PRINCIPLE OF ELECTROKINETIC REMEDIATION

In this technique electric field is created in a soil matrix by applying a low voltage current (DC) to electrode placed in the soil. As a result of electric field, heavy metal contaminants are mobilized, concentrated at electrodes and extracted from the soil. The current density is generally order of mill amperes per square centimeters (mA/cm<sup>2</sup>) or an electric potential difference on the order of few volts per centimeter across the electrodes. When DC current is applied to the electrodes, an electrical field develops between the

anodes and cathodes. The application of the electric field has several effects on the soil, water, and contaminants. The direction and quantity of contaminant movement is influenced by the contaminant concentration, soil type and structure, and the mobility of contaminant ions, as well as the interfacial chemistry and the conductivity of the soil pore water. In low permeable soils, mainly three types of electro kinetic phenomena have been identified which are discussed below. These effects include electro migration, electro osmosis, changes in pH, and electrophoresis changes in pH, and electrophoresis.

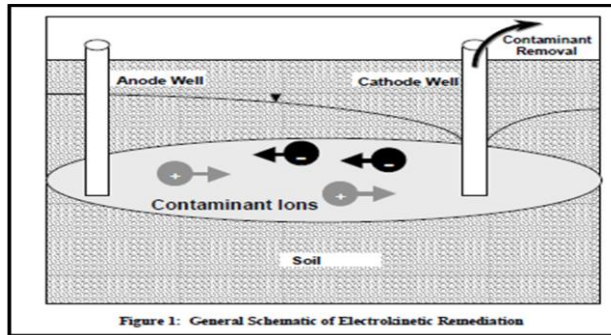


Fig-1 General schematic of Electro kinetic Remediation

Oxidation of water occurs at the anode and generates hydrogen ( $H^+$ ) ions  $H^+$  ions generate an acid front, which migrates to the cathode. In contrast, reduction of water occurs at the cathode and generates hydroxyl ( $OH^-$ ) ions, which migrate as a base front towards the anode. Extraction and removal of the contaminants may be accomplished by electro deposition, precipitation or co-precipitation at the electrode, or removal and treatment of water containing the mobile contaminant species from the electrode wells. There is a lack of understanding of the technology's effects on naturally occurring ions and how these effects impact mobilization and removal of the target contaminants. There is a lack of understanding of the impact that electrode shape and electrode placement will have on the electric field shape and intensity formed within the soil matrix. The soil chemical and biological factors that may limit the application of electrokinetic remediation have yet to be adequately quantified.

#### A. *Electro-osmosis*

When the electric field is applied across the wet soil, the soil water electrolyte behaves as an electrochemical cell. In this electrolytic cell, by convention, the cathode is designated negative and the anode is positive. Current flows from positive anode to negative cathode, opposite to electron flow. Due to the necessity for balancing the negatively charged soil particles, the number of mobile cations accumulates and exceeds the number of anions. Under the electric potential, there is a net movement of cations toward the cathode and the pore fluid is dragged with them toward the cathode compared to the relatively stationary charged soil surface. This flow is termed as electro-osmotic flow.

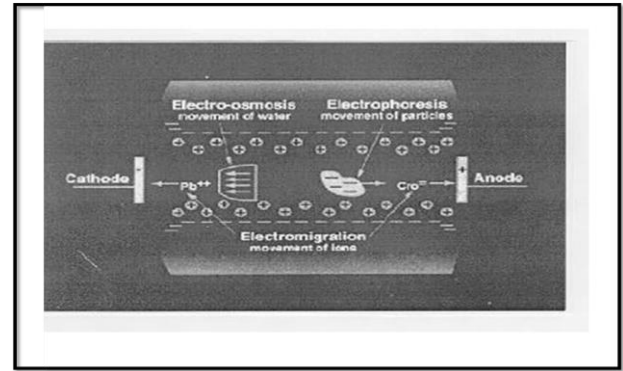


Fig no. 2 Contaminant Transport process

#### B. *Electro-migration*

Electro-migration is the transport of charged ions and ion complexes to the oppositely charged electrode under the influence of an electric field. Electro-migration takes place when highly soluble ionized inorganic species, including metal cations, chlorides, nitrates and phosphates, are present in moist soil environments. The negatively charged anions will move towards the anode and the positively charged cations will move towards the cathode.

#### C. *Electro-phoresis*

When a DC field is applied across the colloidal suspension, charged particles are attracted electro-statically to one of the electrodes and repelled from the other. Negatively charged clay particles move towards anode as shown in Fig 2. This phenomenon is called electrophoresis. The difference between electrophoresis and electro-osmosis is, electrophoresis involves discrete particles transport through water and electro-osmosis involves water transport through a continuous soil particle network.

### IV. AIM AND OBJECTIVES OF THE WORK

As explained earlier, there are many conventional methods which are used for the removal of heavy metals from soil like Excavation, Chemical Reduction, soil washing method, Soil vapor Extraction, Thermal remediation, Electro kinetic Remediation etc. But all these methods are not suitable for the fine grained soil because of its low permeability except Electro kinetic remediation. The specific objectives of present study are:

- To study the effect of various process parameters like, electrode spacing and current density on the extraction rate keeping the soil characteristic constant.

Carry out the study on the naturally occurring soils and that to in the same state as it exists, instead of using kaolinite or artificially creating the contaminated soil mass.

### V. EXPERIMENTAL METHODOLOGY

In order to evaluate the effect of individual parameter on extraction rate Graphite electrode is used. While electrode

spacing is kept 75mm, 100mm and 150mm and current density is kept 37.5v, 50v and 75v. The extraction rate is measured in each trial. It is planned to have three no. of cell at a time in operation and subjecting the soil sample with identical characteristic.

The soil sample for experimental purpose is collected from the Industrial zone of Dahej, Gujarat. In these region lots of chemical and industrial industries exist so the heavy metal is observed in this region. Following are the physical and chemical properties of the soil which will be use our experimental set up. Index properties of soil samples are to be determined before stating the experimental work.

#### A. Experimental Set Up



Fig-3 Experimental Set Up of Electro kinetic Remediation

A schematic test setup used is shown in Fig 3. The test set up consists of cylindrical cell made up of Perspex material containing a sample compartment of 5.0 cm inner diameter and variable length having two electrode compartments, with an anode reservoir and cathode reservoir. At each end of the soil cell, a cork with nylon net is used as reinforcement. It is immediately provided next to filter paper to prevent spreading of fine soil particles from entering into the reservoir. Gas vents were provided in both the electrode compartments to pass gases generated from the electrolysis process. Graphite anode and cathode were fitted at the end of reservoir and were connected with a direct-current power supply. The unit is powered by rectifier working on single phase AC. The voltage, current and various incremental voltages of the unit are monitored and logged manually using digital millimeters at regular intervals fixed on the rectifier panel. The electro kinetic test setup used in this study was similar to that used in previous electro kinetic research by Reddy *et al.* 1997.

#### B. Experimental Process

Approximately 500 gm of oven dry soil passing through 425 micron sieve is used for each test. Potassium Dichromate dissolved in distilled water and mix with the soil. Additional required Distilled water is added as per required moisture content. After mixing the soil sample is allowed to equilibrate for 1 hr. The moist soil is placed into the electrokinetic cell in layers, and each layer was compacted thoroughly using a stainless steel rammer so that the amount

of void space is minimized. Once the cell was full with soil, the sides and the outer part of the cell were completely cleaned. Then, anode and cathode compartments and reservoirs were attached to the cell and the anode reservoir was filled with appropriate purging solution. Each electrode Compartment consisted of a valve to control the flow into the cell, a slotted Graphite electrode and a porous stone. The electrode compartments were connected to either end of the cell using screws. Exit ports were created in the electrode compartments, and the tubing was attached to these ports to allow the gases generated due to the electrolysis of water to escape. The other end of these gas tubes was connected to the reservoirs to collect any liquid that was removed along with the gases. A power source was used to apply a constant voltage to the electrodes, and a multi meter was used to monitor the voltage and measure the current flow through the soil sample during the test.

In the photograph the test cell is made acrylic transparent material of length 75mm,100mm,150mm and 50 mm inside diameter and 60mm outer diameter each. Anode and cathode reservoir are made from Perspex material Teflon. In each compartment Filter paper , porous stone , Filter paper and graphite electrode whole assembly is placed. Rectifier is used to convert the A.C. current into D.C. current. In the rectifier one controller switch is fitted so required amount of voltage can be applied to the set up. Rectifier set up is made in such a way that current is applied on three set up at a time .Current passes from set up can recorded in each individual window. here in photograph Electrical pannel made in such a way that we can perform the 3 experimental in a series at a time



Fig-4 Photograph Of Rectifier

#### C. Soil Digestion Method

At the completion of each test, the reservoirs and the electrode assemblies were disconnected, and the soil specimen was extruded from the cell using a mechanical extruder. The soil specimen was sectioned into four equal parts. Each part was weighed and preserved in a glass bottle and was used to chromium concentrations. From each soil section, 5 g of soil was taken and mixed with 25 ml distilled water and filter it. If sample has been filtered and only chromium is desired, start analysis within 24 h of collection and proceed. Add 5 ml HNO<sub>3</sub> and 10 ml H<sub>2</sub>SO<sub>4</sub>. Boil the sample to the appearance of so<sub>3</sub> fumes. Adjust solution ph 0.7 to 1.3 Transfer solutions to a 100 ml volumetric flask , dilute to 100ml, and mix . Add 2.0 ml Diphenylcarbazide solution mix and let stand 5 to 10 min for full color

development. Some portion of this solution for required metal determination. Transfer an appropriate portion to a 1-cm absorption cell and measure its absorbance at 540 nm. Use distilled water as reference. Correct absorption reading of sample by subtracting absorbance of a blank carried out through the method.

D. Spectrophotometer



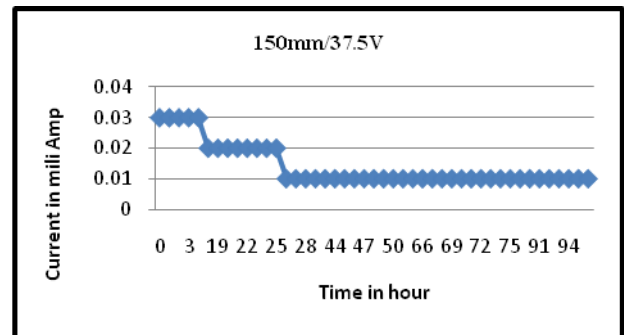
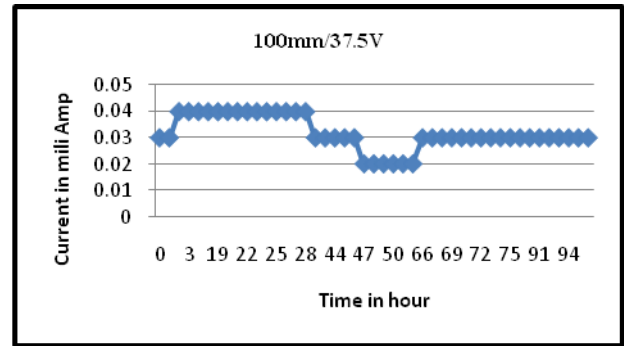
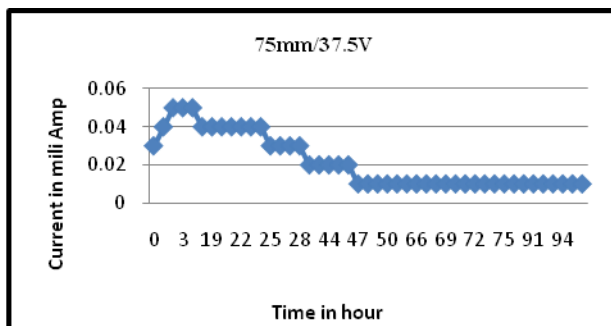
Fig-5 Photograph Of Spectrophotometer

Use of spectrophotometer determines heavy metal like, mercury, cadmium, lead, nickel, copper, zinc, chromium. For metal determination visible spectro photometer is used. Light rays are passed through sample and according to metal concentration and color light is absorb in the sample. For this purpose used colorimetric method for chromium removal. (APHA standards)

VI. RESULT AND ANALYSIS

After completed the soil sample preparation, soil sample filled in the different length of cell and different current density is applied at 96 hour. During the experimental process note down the current density, extraction rate, purging solution, at different time.

Graph for Time Vs Current density for Exp



A. Chromium Removal from the soil

Use different length of cell and different current density chromium can be removal from the soil. Following are the result of chromium removal of different length of cell at different voltage

Table no-5 Removal of chromium

Test	Length of cell	Voltage	% Removal
1	75 mm	37.5	70.35 %
2	100mm	37.5	68.49 %
3	150mm	37.5	78.49%
4	75mm	50	29.03 %
5	100mm	50	62.83 %
6	150mm	50	60.45 %
7	75mm	75	70. 81 %
8	100mm	75	78.43 %
9	150mm	75	74.7 %

CONCLUSION AND LIMITATIONS

A. Conclusion

- The percentage removal increase with increase in length of the cell. However at higher potential percentage removal is maximum for 100 mm cell length when at lower potential the percentage removal is minimum at 100mm cell length.
- From the analysis it can be observed percentage removal increase with increase in spacing.

- At minimum spacing unbalance charge is minimum compare to that at higher spacing however general believe that with increase spacing percentage removal is increase.
- The range of spacing selected should be widening, higher spacing through high unbalancing charge resulting in higher removal but again same result into practically no removal in the mid portion of the cell. Where in influencing of the electrode is minimum.
- Increase current density increases the rate of removal.
- With the same spacing and current density total removal can be higher if process is extended for longer time.
- The current density & spacing are complementary variables for optimization of removal.

#### B. *Limitations*

- In this experiment work only CH type soil is used.
- Removal efficiency of different heavy metals can be determined. Acidic condition and electrolytic decay corrode some anode material.
- Removal efficiency of different heavy metals can be determined.
- The experiments conducted in this work have been one dimensional.

#### REFERENCES

- [1] Krishna Reddy ,”Electro Chemical Remediation Technology for polluted soils “
- [2] Mahesh Rawat ,”Electro kinetic Remediation of Cadmium-Contaminated Soil Using Zero-Valent Iron Particles”
- [3] Dr. Kevin Gardner ,”Electrochemical Remediation and Stabilization of Contaminated Sediments “
- [4] Ashraf Al-Hamdani Krishna Reddy, “Electrokinetic Remediation Modeling Incorporating Geochemical Effects”
- [5] Mary Page and Christopher Page,”Electro remediation of Contaminated Soil”
- [6] Enhanced electrokinetic extraction of heavy metals from soils assisted by ion exchange membranes: By Won-Seok Kim, Soon-Oh Kim, Kyoung-Woong Kim
- [7] Removal of cadmium using electro kinetics-effect of enhancement agents:By Srivastava, R.K., Tiwari, R.P. and Bala Ramudu P
- [8] Enzo Lombi, Walter W. Wenzel, and Domy C. Adriano,”Soil contamination, risk reduction and remediation”
- [9] S.Liu and H. Paul Wang ,”In Situ Speciation Studies of Copper–Humic Substances in a Contaminated Soil during Electro kinetic Remediation”

# Feasibility Study of Compartmental Elevated Water Tank as Per IS: 1893 Draft Code (Part-2)

A. Chirag N. Patel, B. H. S. Patel

**Abstract**— This paper presents the dynamic analysis of a rectangular and circular elevated liquid storage tank with multiple compartments. Geometrical consideration for tank container has been taken as per standard height to length/diameter ratio and complete finite element models are to be prepared and analyse using SAP 2000 software. The water mass inside the tank is distributed appropriately to take into account the hydrodynamic forces developed due to impulsive as well as convective effects. Base shear and overturning moment results are compared with the values which are calculated from the expressions given in draft code IS 1893(part-2). Moreover it can be seen from the analysis results that above parameters are increases with number of compartments in both types of compartmental tanks. A conclusion from the design showed that, compartmental tanks of circular shape is more economical than the rectangular one and find considerable difference in the cost of staging part of rectangular tank while considered three compartments.

**Index Terms**— Compartmental tank, staging, water tank.

## I. INTRODUCTION

Water tanks are very important structure for public utility and industry having basic purpose of to secure constant water supply at the longer distance with sufficient static head under the effect of gravitational force. The overhead tanks which have been the inevitable part of water supply system, by help of which the required water head can easily be achieved to erecting it on purpose made supporting towers and water can be made available to all by the action of gravity. It is also essential to ensure that, requirements such as water supply is not hampered during an earthquake and should remain functional in the post-earthquake period. In such situations the elevated tanks may prove most handy tool for the purpose of water distribution and fire protection. Earthquakes are the most destructive among all the natural hazards. Elevated water tank which consist of a huge mass supported on the top of a slender staging are particularly susceptible to earthquake damage. So, water tank structure requires proper attention in seismic analysis, design and construction as well. The dynamic response of liquid storage tanks subjected to earthquake has been a subject of numerous studies in the past 30 years. In current Indian code for seismic design, limited provisions on

seismic design of elevated tanks are given. Those provisions are highly inadequate compared to present international practice. Most of the previous studies were focused on the tank containing liquid considering only one mass and it does not cover important aspects for analysis and design of water tanks related to the effect of hydrodynamic pressure of the water, which produce due to vibration of the tank when earthquake strikes. But, After the Bhuj earthquake, revision of current Indian code became inevitable. Hence it was decided to develop guidelines [1] under the project “Review of Building Codes and Preparation of Commentary and Handbooks” assigned by the Gujarat State Disaster Management Authority (GSDMA), Gandhinagar to the Indian Institute of Technology Kanpur in 2003. The draft code for liquid retaining structures is one of the outcomes of the project. The present study is based on this draft code.

Looking today’s scenario of infrastructure growth, there is need arising to have a more than one storage tank also for general purpose. A compartmental tank contains more than one compartment in container itself. It facilitates to store more than two types of liquid in one tank i.e. potable and domestic water can be supplied from the different compartments of the same elevated tank or to make extra storage for continuous supply during maintenance work and even when it is being cleaned. So, instead of constructing two or more storage tanks, it is quite economical to construct large size storage tank with two or multi compartments. But, while using water tank with compartment more than one there is a problem of eccentric loading arrives in the case of one container empty, and due to that effect of torsion has been take place. There are no clear guidelines provided by draft code for the consideration of hydrodynamic effect on compartmental liquid storage tanks. In this context, the present study is an effort to identify the addition on the seismic parameters and construction cost considering compartmental storage tank.

## II. OVERVIEW OF DRAFT CODE IS 1893 (PART – II)

Draft code IS: 1893 (Part-II) [3] has been prepared in accordance with generally recognized engineering principles and practices, and also with the reference of many international codes, standards and guidelines. Also, seismic analysis of water tanks considering hydrodynamic effect has been covered and explained in detail. Following important provisions and changes have been incorporated as compared to that of provisions of previous version of IS 1893:1984[4].

- Analysis of ground supported tanks.

A. Ph.D. Candidate from Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India. (e-mail: cnpatel.693@gmail.com)

B. Associate Professor of Applied Mechanics Department, L. D. College of Engineering, Ahmedabad, Gujarat, India. (e-mail: dr.hspatel@yahoo.com)

- For elevated tanks, the single degree of freedom is replaced by two degree of freedom idealization.
- Bracing beam flexibility is explicitly included in calculation of lateral stiffness of tank staging.
- The effect of convective and impulsive hydrodynamic pressure distribution in the analysis.
- Effect of vertical ground acceleration on hydrodynamic pressure and pressure due to wall inertia.
- Sloshing effect of water and maximum sloshing wave height.
- P-Delta effect for elevated water tank.

Hydrodynamic effect is considered by dividing water into two different masses, namely impulsive and convective. When the tank containing liquid with free surface is subjected to horizontal earthquake ground motion, tank wall and liquid are subjected to horizontal acceleration. The liquid in the lower region behaves like a mass that is rigidly connected to the tank wall. This mass is termed as impulsive liquid mass, which, accelerates along with the wall and induces impulsive hydrodynamic pressure on tank wall and on base. The liquid mass in the upper region of tank undergoes sloshing motion. This mass is termed as convective liquid mass and exerts convective hydrodynamic pressure on wall and base. For representing these two masses and in order to include the effect of their hydrodynamic pressure in analysis, spring mass model is adopted for ground-supported tanks and two-mass model for elevated tanks.

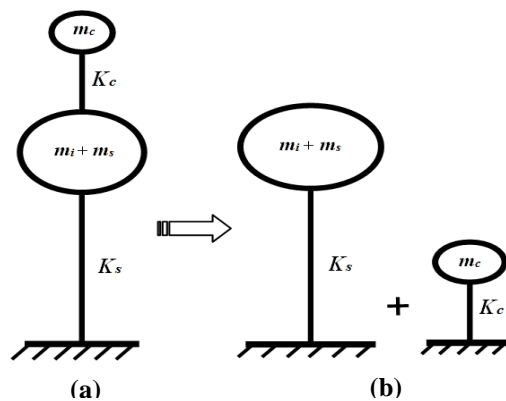


Fig.1: Two mass model for elevated tank

In spring mass model convective mass ( $m_c$ ) is attached to the tank wall by the spring having stiffness ( $K_c$ ), whereas impulsive mass ( $m_i$ ) is rigidly attached to tank wall. For elevated tanks two-mass model is considered, which consists of two degrees of freedom system. Spring mass model can also be applied on elevated tanks, but two-mass model idealization is closer to reality. The two-mass model is shown in Fig.1. where,  $m_i$ ,  $m_c$ ,  $K_c$ ,  $h_i$ ,  $h_c$ ,  $h_s$ , etc. are the parameters of spring mass model and charts as well as empirical formulae are given for finding their values. The parameters of this model depend on geometry of the tank and its flexibility. For elevated tanks, if the shape is other than circular or rectangular, then the values of spring mass parameters can be obtained by

considering an equivalent circular tank having same capacity with diameter equal to that of diameter at top level of liquid in original tank.

The two-mass model was first proposed by G. M. Housner [5] and is being commonly used in most of the international codes. The response of the two-degree of freedom system can be obtained by elementary structural dynamics. However, for most of elevated tanks it is observed that both the time periods are well separated. Hence, the two-mass idealization can be treated as two uncoupled single degree of freedom system as shown in Fig.1 (b). The stiffness ( $K_s$ ) is lateral stiffness of staging. The mass ( $m_s$ ) is the structural mass and shall comprise of mass of tank container and one-third mass of staging as staging will acts like a lateral spring. Mass of container comprises of roof slab, container wall, gallery if any, floor slab, floor beams, ring beam, circular girder, and domes if provided.

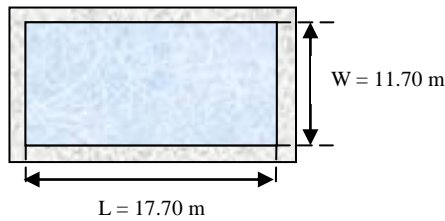
### III. MODELING OF WATER TANKS

A reinforced concrete elevated water tank of storage capacity 1000 m<sup>3</sup>, with rectangular and circular type of container shape is selected to form the compartmental tanks considering one, two and three number of compartments. Moreover, there is no constructional difficulty as far as the formwork availability is concerned with this kind of shape. Analysis and design of this type of container is also easy in comparison with other types of container. To perform the seismic analysis of compartmental water tank, dimensions like seismic weight, time period, horizontal acceleration coefficient, base shear, etc. has been decided to evaluate Initially, sizes of different components are assumed in proportion with the size of container as shown in Fig. 2 & 3.

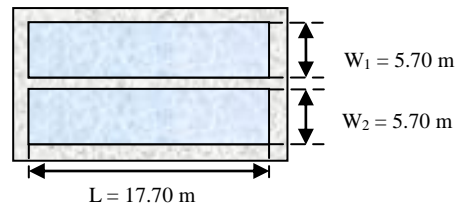
On the bases of required volume of the tank, dimensions are found out first in terms of the length & width for rectangular tank and diameter for circular tank with considering appropriate height. Here generally height of tank is taken near about 0.30 times length or diameter. Also for rectangular tank width of tank is taken about 0.65 times length. The dimensions of other components are also proportioned. When a single compartment tank is to be converted in multi compartmental tank, care should be taken to produce compartments of required volume and more or less equally distributed in terms of total volume of tank. Also, design of the container has been carried out for hydrodynamic pressure.

The other assumed data are: Shape of container : Rectangular & Circular

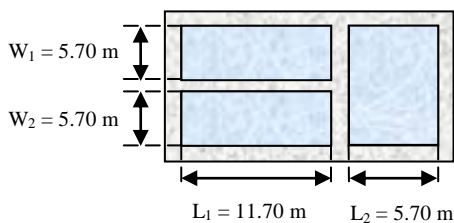
- Capacity of tank : 1000 m<sup>3</sup>
- Height of container : 5.2 m
- Freeboard = 0.3 m
- Type / Height of staging : Trestle / 15 m
- Shape of column : Circular
- Depth of foundation : 3.0 m
- Soil Type : Soft Soil
- Grade of concrete and steel: M 25 and Fe 415



(a) Single Compartment Rectangular Tank

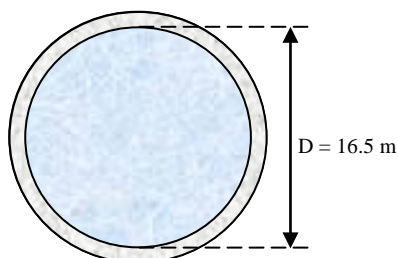


(b) Two Compartment Rectangular Tank

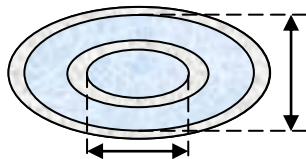


(c) Three Compartment Rectangular Tank

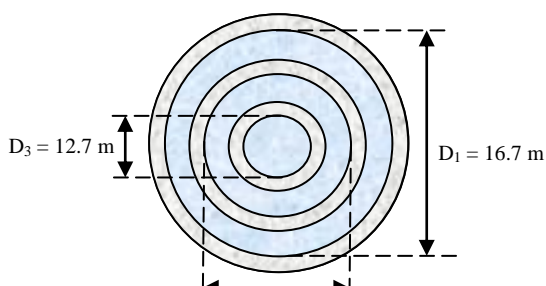
Fig. 2: Rectangular Tank with different number of compartment



(a) Single Compartmental Circular Tank (1000 m<sup>3</sup>)



(b) Two Compartmental Circular Tank (1000 m<sup>3</sup>)

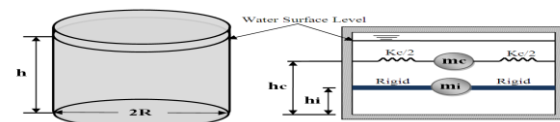


(c) Three Compartmental Circular Tank (1000 m<sup>3</sup>)

Fig. 3: Circular Tank with different number of compartment

Models are prepared separately for impulsive and convective masses. The concept that enables analysis of elevated water tanks as a single lumped-mass model was suggested in the 1950s (Chandrasekaran and Krishna, 1954). The Indian seismic code IS: 1893-1984, requires elevated tanks to be analyzed as a single degree of freedom system i.e. a one-mass system, which suggests that all fluid mass participates in the impulsive mode of vibration and moves with the container wall. It must be stated that this can be a realistic assumption for long and slender tank containers with a height-to-radius ratio exceeding four. Also, the ACI 371R-98 (1995) suggests that the single lumped mass model should be used when the water load ( $W_w$ ) is 80% or more of the total gravity load ( $W_g$ ) that includes: the total dead load above the base, water load and a minimum of 25% of the floor live load in areas that are used for storage. Two significant points should be discussed for this concept. The first point is related to the behavior of the fluid. If the container is completely full of water, this prevents the vertical motion of water sloshing, so the elevated tank may be treated as a single-degree-of-freedom system in such a case. But most elevated tanks are never completely filled with liquid. So, when the fluid in the container (vessel) oscillates, this concept fails to characterize the real behavior. The other point is related to the supporting structures. As the ductility and the energy-absorbing capacities are mainly regulated by the supporting structure, this is important for the seismic design of elevated tanks. The elevated tanks can have different types of supporting structures, which could be in the form of steel- frame, a reinforced concrete shell, a reinforced concrete frame or a masonry pedestal.

The equivalent spring-mass models have been proposed by some researchers



and Indian seismic draft code, to consider the dynamic behavior of the fluid inside a container, suggests that elevated tanks to be analyzed considering two-mass idealization of the tank. The equivalent spring-mass models have been proposed by code is shown in Fig. 4. The fluid is replaced by an impulsive mass ( $m_i$ ) that is rigidly attached to the tank container wall and by the convective mass ( $m_c$ ) that is connected to the walls through the springs of stiffness ( $K_c$ ). According to the literature survey, although only the first



convective mass may be considered (Housner, 1963), A single convective mass is generally used for the practical design of the elevated tanks (Haroun and Housner, 1981; Livaoglu and Dogangun, 2005) and higher modes of sloshing have negligible influence on the forces exerted on the container. As practical analyses are presented in this study, only one convective mass is taken into consideration.

Fig. 4: Spring mass analogy for tank container

Where,

- $m_c$  = Convective mass
- $m_i$  = Impulsive mass
- $h$  = Depth of fluid
- $h_c$  = Height of Convective mass
- $h_i$  = Height of impulsive mass
- $R$  = Inner radius of container
- $K_c$  = Stiffness of the convective mass spring

The dynamic characteristics are estimated by using the expressions provided by code. In this table,  $m$  is the total mass of the fluid. Similar equivalent impulsive ( $m_i$ ) and convective ( $m_c$ ) masses and their heights are calculated, respectively, for the overturning moment. Impulsive and convective masses are applied at respective heights for all the compartments of tank with proper rigid and spring stiffness as shown in Fig. 5. For the convective mass a link with the stiffness of  $K_c$  has been used.

IV. SEISMIC ANALYSIS OF COMPARTMENTAL TANK

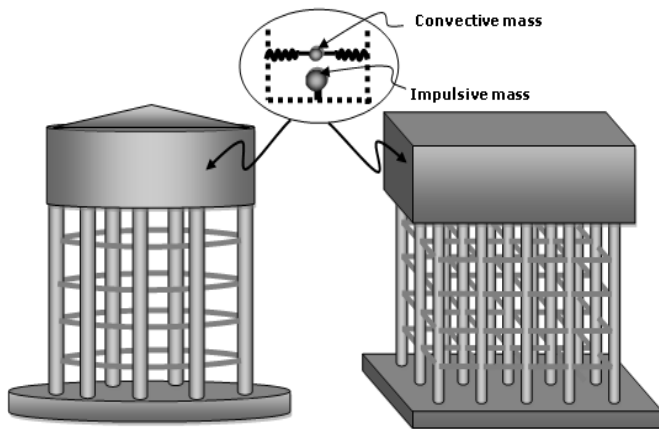


Fig. 5: Tank models with different mass application

Tank has a frame supporting structure in which columns are connected by the beams (horizontal bracings) at regular intervals, at 3, 6, 9 and 12 m height.

For the analysis, some additional data required is as follows,

- Type of soil = Soft Soil
- Importance of structure = Tank used for water storage
- Type of Moment Frame = S.M.R.F.
- Width of brace = 300 mm
- Depth of brace = 500 mm

- Diameter of column = 350 mm

Total twelve FEM models are prepared with the use of SAP 2000 software considering fixed base condition. The damping values for the reinforced concrete elevated tanks are taken as 5% for the impulsive mode and 0.5% for the convective mode, as recommended in most literature. Models are prepared separately for impulsive and convective masses, Out of twelve models six for each rectangular and circular tank. Impulsive and Convective masses are applied at respective heights for all the compartments of tank as shown in Fig 5. For the convective mass a link with the stiffness of  $K_c$  has been used.

In output result, various quantities of seismic analysis like parameters of spring mass model, impulsive mass, convective mass, lateral stiffness of staging, empty weight of container, impulsive and convective time period, design horizontal seismic coefficient, base shear, and base moment for both the modes are obtained as per draft code calculations and from software models. Value of base shear ( $V$ ) and overturning moment ( $M$ ) at base of rectangular and circular compartmental tanks has been compared as shown in Table 1, 2, 3 & 4 respectively.

TABLE 1  
BASE SHEAR FOR RECTANGULAR TANK

Rectangular Tank (1000 m <sup>3</sup> )	Draft code		SAP 2000	
	Base Shear (V) kN	% Increase in Base Shear (V)	Base Shear (V) kN	% Increase in Base Shear (V)
	890.62	-	912.47	-
	909.44	2.06	932.72	2.17
	993.38	10.52	1041.89	12.64

TABLE 2  
BASE SHEAR FOR CIRCULAR TANK

Circular Tank (1000 m <sup>3</sup> )	Draft code		SAP 2000	
	Base Shear (V) kN	% Increase in Base Shear (V)	Base Shear (V) kN	% Increase in Base Shear (V)
	818.14	-	820.55	-
	866.81	5.61	877.73	6.51
	911.23	10.48	908.04	9.84

TABLE 3  
OVERTURNING MOMENT FOR RECTANGULAR TANK

Rectangular Tank (1000 m <sup>3</sup> )	Draft code		SAP 2000	
	Overturning Moment (M) kNm	% Increase in Overturning Moment (M)	Overturning Moment (M) kNm	% Increase in Overturning Moment (M)
	16995.33	-	17102.08	-






	17283.78	1.67	17433.98	1.90
	19195.92	11.63	19724.45	13.52

TABLE 4  
OVERTURNING MOMENT FOR CIRCULAR TANK

Circular Tank (1000 m <sup>3</sup> )	Draft code		SAP 2000	
	Overturning Moment (M) kNm	% Increase in Overturning Moment (M)	Overturning Moment (M) kNm	% Increase in Overturning Moment (M)
	15091.38	-	15323.12	-
	16340.75	7.65	16786.92	8.72
	17204.92	12.67	17978.25	15.35

V. COST ANALYSIS








Further study has been carried out by considering different number of staging columns for same capacity (1000 m<sup>3</sup>) of rectangular and circular compartmental water tank. The circular columns are used with the size of 350 mm diameter for all the cases. During the analysis of the Rectangular compartmental tank number of columns has been taken as 35 and then reduced to 29. The effect of torsion plays important role for the rectangular compartmental tank. Also, the same

study is carried out for circular compartmental tank with number of columns 25 and then reduced to 19. So, quantity of steel and concrete are calculated for all the cases of tank compartment full and empty conditions as shown in Table 5 & 6. Also the quantity variation of steel and concrete is presented graphically in Fig. 6, 7, 8 & 9. From the analysis and design, column with maximum % of steel required has been chosen for calculating the total quantity of steel. Concrete quantity is calculated from the cross sectional area of the staging sections.

CONSLUSIONS

After the comparison of numerical values available in Table 1, 2, 3 & 4 base shear and moment in Rectangular and Circular Compartmental Tank slightly increases with number of compartments, but it is more due to additional internal wall which is provided to make compartment. The impulsive mass would increase; it would also contribute in an increment of base shear and base moment. Also, not much difference has been observed in base shear from model analysis compared to calculated codal value. Base shear and base moment in Rectangular and Circular Compartmental Tank, increases with number of compartments. Therefore, provision of compartment in tank provides intended purpose during life span without affecting much difference in economy.

TABLE 5  
QUANTITY CALCULATION FOR RECTANGULAR TANKS WITH COST COMPARISON

Container	Water In Tank	Steel Area mm <sup>2</sup>	% Steel	Total Steel Quantity Kg	Total Concrete Quantity Kg	Total Cost of Staging Rs.	Steel Area mm <sup>2</sup>	% Steel	Total Steel Quantity Kg	Total Concrete Quantity Kg	Total Cost of Staging Rs.	Remark
		35 No. Column of 350 mm Dia.					29 No. Column of 350 mm Dia.					
One Compartment		832	0.86	20894.32	136.65	1092233	1319	1.37	24654.10	127.37	1183778	No Torsion
Two Compartment		959	1.00	21902.21	136.52	1124083	1589	1.65	25637.55	127.24	1214846	Torsion
		964	1.00	21924.19	136.82	1125716	1476	1.53	25225.96	127.76	1203287	No Torsion
Three Compartment		1197	1.24	22948.46	136.39	1157160	1768	1.84	26289.54	127.16	1235461	Torsion
		4178	4.34	36052.93	134.72	1571326	5324	5.53	39241.91	125.51	1644822	Torsion
		3964	4.12	35112.19	134.84	1541594	4987	5.18	38014.42	125.67	1606038	Torsion
		1106	1.15	22548.42	137.20	1146869	1676	1.74	25954.44	127.98	1227280	No Torsion

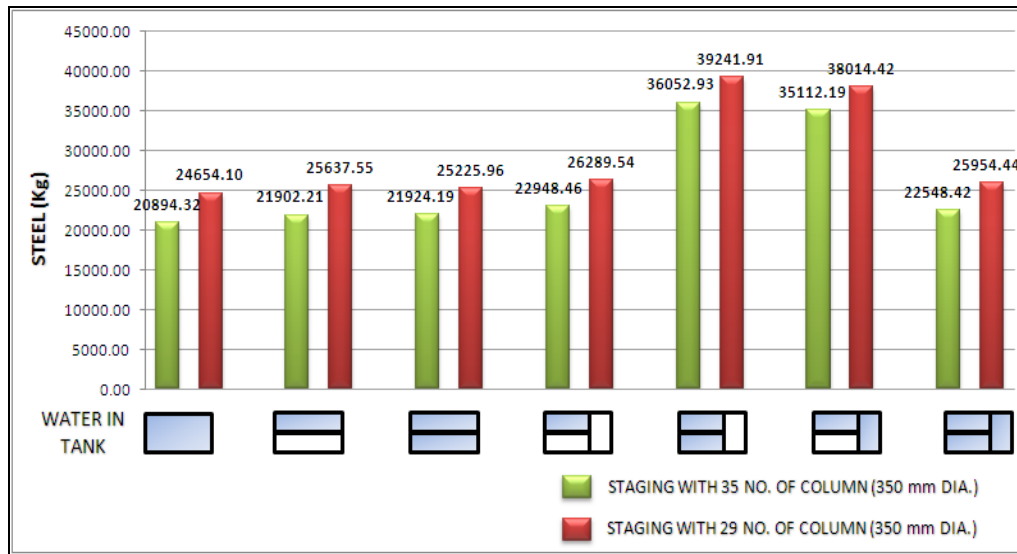


Fig.6: Steel quantity of staging for Rectangular Tanks

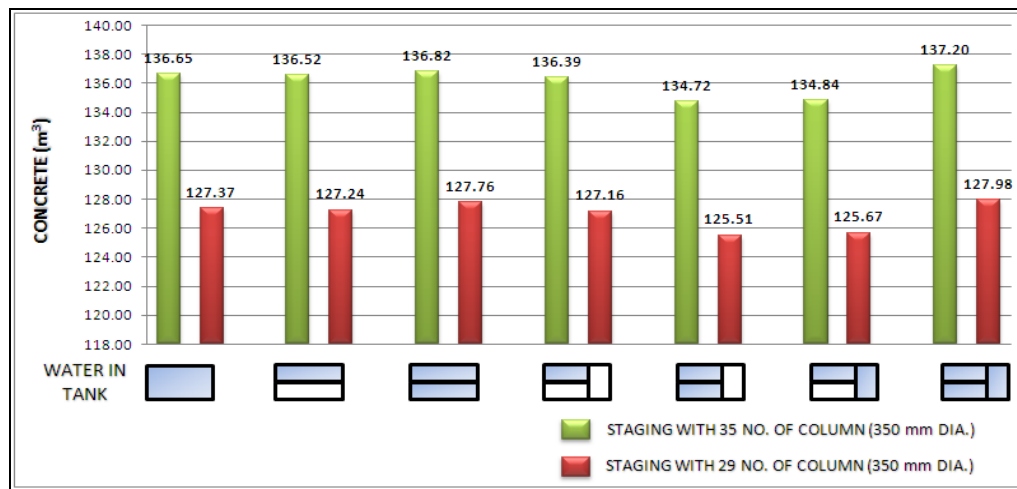


Fig.7: Concrete quantity of staging for Rectangular Tanks

TABLE 6  
QUANTITY CALCULATION FOR CIRCULAR TANKS WITH COST COMPARISON

Container	Ring of Column	Steel Area mm <sup>2</sup>	% Steel	Total Steel Quantity Kg	Total Concrete Quantity Kg	Total Cost of Staging Rs.	Steel Area mm <sup>2</sup>	% Steel	Total Steel Quantity Kg	Total Concrete Quantity Kg	Total Cost of Staging Rs.
	25 No. Column of 350 mm Dia.						19 No. Column of 350 mm Dia.				
One Compartment	Outer	1129	1.17	27542.68	140.31	1316327	2207	2.29	31217.79	132.19	1408758
	Middle	1002	1.04				1927	2.00			
	Inner	909	0.94				1792	1.86			
Two Compartment	Outer	1276	1.33	27999.46	132.90	1307973	2429	2.52	31991.22	124.73	1410382
	Middle	1158	1.20				2165	2.25			
	Inner	1072	1.11				1961	2.04			
Three Compartment	Outer	1400	1.46	28505.62	140.19	1346769	2677	2.78	32328.25	132.05	1443859
	Middle	1337	1.39				2432	2.53			
	Inner	1223	1.27				2213	2.30			

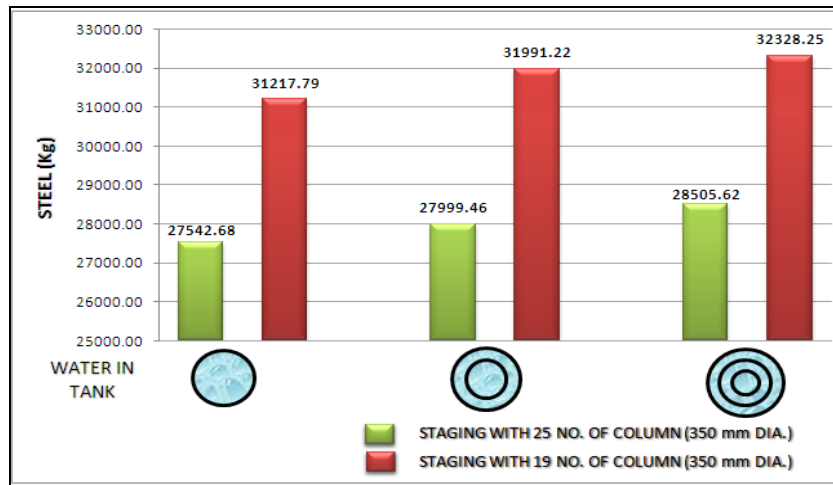


Fig.8: Steel quantity of staging for Circular Tanks

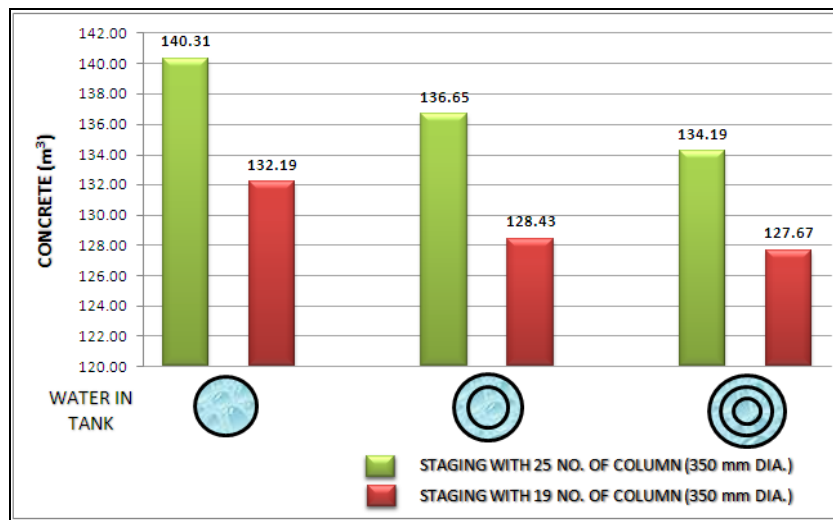


Fig.9: Concrete quantity of staging for Circular Tanks

The quantity of steel for staging has been calculated from the column considering maximum % of steel under worst loading condition for all the above cases. From the observation, the steel quantity for the rectangular compartmental tank is increased about 15 % for different number of column configuration. And for compartmental wise comparison, it is increased from compartment one to two and again two to three about 13% and 40% respectively. Concrete quantity is not much changed compared to steel quantity. The steel quantity for the circular compartmental tank is increased about 12 % for different number of column configuration. And for compartmental wise comparison difference is not more than 4% because of negligible torsional effect as compared to rectangular compartmental tank.

However, from the design it is concluded that, for compartmental tanks circular shape is more economical than the rectangular one. Due to the torsional effects in rectangular compartmental tank it requires more steel compared to circular one. And from the cost comparison of

staging it is seen that, for circular tanks little difference has been observed. But for rectangular tank especially for three compartments difference in cost is considerable.

REFERENCES

- [1] IITK-GSDMA Guidelines for Seismic Design of Liquid Storage Tanks Provisions with commentary and explanatory examples.
- [2] IS 11682:1985, “Criteria for Design of RCC Staging for Overhead Water Tanks”, Bureau of Indian Standards, New Delhi.
- [3] Draft IS: 1893 (Part-II, Liquid Retaining Tanks) Criteria for Earthquake Resistant Design of Structures, Bureau of Indian standards, New Delhi, India.
- [4] IS: 1893-1984 (part-2), “Criteria for Earthquake Resistant Design of Structures”, Bureau of Indian Standards, New Delhi.
- [5] George W. Housner, “The dynamic behaviour of water tanks”, *Bulletin of the Seismological Society of America*, February-1963, Vol-53, No.2, Pg. 381-387.

# Former Failure Assessments of RC Elevated Water Tanks: Literature Review

A. Chirag N. Patel, B. H. S. Patel

**Abstract**— This paper provides a literature review on the assessment regarding causes for early deterioration and vulnerability of elevated water tanks as specially staging part under lateral force due to an earthquake or tsunami using a systematic in situ documentation, survey and empirical damage scale. As, due to failure of supporting systems, study considered strength analysis of a few damaged shaft and frame type staging which clearly shows them deficient due to underestimation of the forces and no suitable provisions delivered for ductile detailing of tank stagings by Indian standards. It highlights the weaknesses of the current Indian practice of seismic design considering factors that affects the ductility of reinforced concrete, and put such special attention towards selection of staging configuration, analysis and design to make elevated water tank structure live at the time of earthquake hazards. Also, discuss provisions for retrofitting or up gradation for future earthquake. Additionally, provide recommendations for life cycle costing, useful life and duration of the completion of the project.

**Index Terms**—Damage assessment, Deterioration, Failure, Water tank.

## I. INTRODUCTION

Elevated water storage tank features to look for are strength and durability, and of course leakages can be avoided by identifying good construction practices. But in reality these structures do not often last as long as they are designed for. In general, water retaining structure distress has been observed very early even in 9 to 10 years of service life due to some problems related to structural aspects and over emphasis of seismic analysis in earthquake prone zones. During the past earthquakes, tanks have suffered with varying degree of damages, which include: Buckling of ground supported slender tanks (Malhotra, 1997), rupture of steel tank shell at the location of joints with pipes, collapse of supporting tower of elevated tanks (Manos and clough, 1983, Rai, 2002), cracks in the ground supported RC tanks, etc.

Water tanks can experience distress due to several reasons such as improper structural configuration design, inferior materials, inferior workmanship, corrosion of reinforcement, wind forces, earthquake forces etc. Due to large mass, especially when it is full, earthquake forces almost govern the

lateral force design criteria for these structures in zones of high seismic activity. In the extreme case, total collapse of tank shall be avoided. However, some damage (repairable) may be acceptable during severe shaking not affecting the functionality of tank. Whatever may be the cause of distress but water tanks should fulfill the purpose for which it has been designed and constructed with minimum maintenance throughout its intended life. This necessitates the overhead tanks to be designed safe to the required degree against all possible forces expected to be encountered during its life time economically. Also, the poor performance of such critical facilities during past seismic events in the country required careful scrutiny of their designs, construction quality and their post-earthquake performance.

TABLE I  
SUMMARY OF REVIEW CONDUCTED ON FORMER FAILURE ASSESSMENTS OF  
ELEVATED RC WATER TANKS

Sr. No.	Authors	Year	Topic
A.	Durgesh C. Rai	2002	Seismic Retrofitting of shaft supported elevated tank
B.	Durgesh C. Rai	2003	Performance of elevated tank in Bhuj Earthquake
C.	Sudhir Singh Bhadauria and Dr. Mahesh Chandra Gupta	2006	Durability performance of water tanks
D.	Hemant B. Kaushik and Sudhir K. Jain	2007	Impact of Tsunami on structures
E.	Sudhir Singh Bhadauria and Dr. Mahesh Chandra Gupta	2007	Testing of deteriorating water tanks for durability assessment
F.	Amjad Masood, Tazyeen Ahmad, Muhammad Arif and V. P. Mital	2008	Failure of water tank in state of Uttar Pradesh

## II. REVIEW OF LITERATURES

### A. Durgesh C. Rai (2002)

This paper discuss unfavorable features of shaft supported elevated tanks in high seismic areas, as evidence with failure of two water tanks in the 1997 Jabalpur earthquake and a great many in the 2001 Bhuj earthquake. Also, identifies the seismic deficiencies of the shaft supports and how they can be retrofitted or upgraded for future earthquakes. It also raised the issue related to the weaknesses of the current Indian practice of seismic design and analysis of structures by the failed support structure of two 10 to 12 year old, 0.5-million-gallon (2270 m<sup>3</sup>) capacity, elevated water tanks in the

A. Ph.D. Candidate from Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India. (e-mail: cnpatel.693@gmail.com)

B. Associate Professor of Applied Mechanics Department, L. D. College of Engineering, Ahmedabad, Gujarat, India. (e-mail: dr.hspatel@yahoo.com)

Jabalpur earthquake of 22 May 1997. The cylindrical shaft-type staging developed circumferential flexural-tension cracks near the base. Similar damages to support structures had been observed in past earthquakes including Bhuj earthquake of 26 January 2001, as shown in Fig. 1.

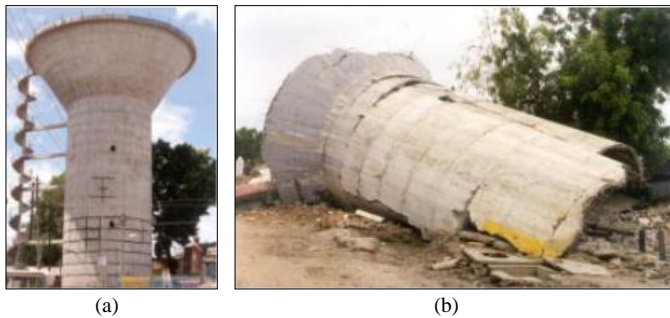


Fig. 1. Water tanks affected during Bhuj earthquake in Morbi (a) and Chobari (b)

Damage observed in tanks supported on shafts where cracks are formed near the base. Some diagonal cracks of shear-flexure origin and some around corners of the window openings were also observed. The flexure-tension cracks in shafts appeared at the level of the first ‘lift,’ a plane of weakness, at 1.4 m above the ground level. These tanks were founded on the compaction-type bored under-reamed piles, and no visible distress to surrounding soil or foundation was noted. The observations indicate that, Thin-walled, circular; shaft supports for elevated tanks behave in a brittle manner at the flexural strength with very small ductility. Also, the Indian seismic code, IS:1893-1984 [7], specifies design forces equivalent to building framing systems and ignores Housner’s two-mass idealization as well as the fact that shaft supports lack the redundancy, damping, and additional strength of building framing systems. Moreover it provides suggestions regarding improvement of ductility using concrete jacketing.

*B. Durgesh C. Rai (2003)*

Discussion regarding, the current designs of supporting structures of elevated water tanks is extremely vulnerable under lateral forces due to an earthquake considering illustration of the Bhuj earthquake is enclosed in this document. During this earthquake many water tank staging of shaft and frame type suffered heavy damage and a few collapsed due to poor ductility of thin shell sections besides low redundancy as well as toughness and weak members with poor brace-column joints respectively in the area of a radius of approximately 125 km from the epicenter (USGS). It also covers the tanks supported on cylindrical shaft type staging which developed circumferential flexural cracks near the base and RC framed type staging located in regions of the highest intensity of shaking collapsed while a few developed cracking near brace-column joint regions. The flexure cracks in stagings were observed from the level of the first lift to several lifts reaching one-third the height of the staging (Fig. 2 (a)), mostly in a circumferential direction and cover the

entire perimeter of the shaft. They usually appear near the edges of the form used during casting of the shaft, which appear to form planes of weaknesses along the shaft’s length, which pass through the thin section and are clearly visible from inside too (Fig. 2 (b)). Frame type tank stagings are also observed collapsed due to not meeting the ductility and toughness requirements for earthquake resistance by brace and column member joints (Fig. 3 & 4).

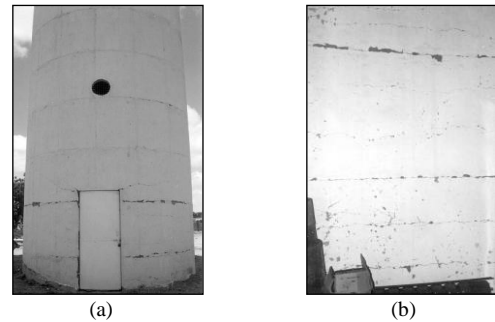


Fig. 2. Water tank which developed circumferential cracks up to one-third height of the staging which are ‘through’ the shell thickness (a) also seen from inside the shaft of Tank (b).



Fig. 3. Collapsed slender and weak framed stagings of water tanks in Manfera village.

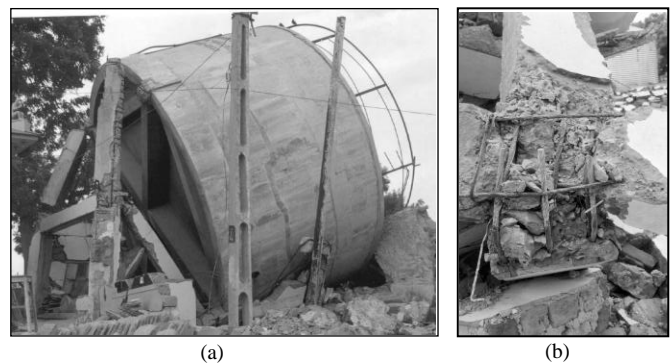


Fig. 4. (a) Severe damage to similar staging required that the water tank be pulled down in Bhachau. (b) Poor detailing of column-brace joints for Manfera tank.

From the strength analysis results of a few damaged shaft type staging which clearly shows that all of them either met or exceeded the strength requirements of IS:1893-1984,

however, they were all found deficient in comparison with requirements of other International Building Code. The provided lateral strength against tensile cracking of staging was either equal or larger than the code required strength and maximum over strength being as large as 170%. In other words, the stagings do meet or exceed the strength requirements of IS:1893-1984 and considered as seismically deficient due to inadequate lateral strength capacity by International Building Code (IBC 2000) under similar seismic exposure conditions.

The supporting structure, especially the framed stagings may look like that used in the building-like structures but its behaviour under seismic loads is very different. Also, lack of redundancy is extremely serious in circular shaft type staging where lateral stability of the structure depends on only a single element, i.e., shaft, and failure of which would severely endanger the lateral stability of the entire structure. Also thin sections of shaft type staging do not have an appreciable level of ductility like frame type staging which can be taken advantage of in dissipating seismic energy and consequently reducing design forces.

For the above reasons, it is been concluded that, Currently IS:1893-1984 underestimates the forces at least a factor of 3 for water tanks. The slender staging that results from the low design forces is a very unfavorable feature for seismic areas. Furthermore, it has been found that circular thin RC sections with high axial load behave in a brittle manner at the flexural strength and, therefore, should be avoided. In comparison, frame stagings of water tanks can be detailed according to provisions of IS:11682-1985 [8] and IS:13920-1993 [9] which refers to the ductility requirements of IS:4326-1976 [10]. The failures of framed staging in epicentral tract was primarily due to non-compliance of ductility provisions of IS codes intended for earthquake resistance, in addition to slender and weak frame members resulting from low seismic design forces.

### C. Sudhir and Mahesh (2006)

Durability is defined as “the structure is able to exist for a long time without significant deterioration and maintaining its level of reliability and serviceability during its lifetime” and other side rapidly increasing rate of deterioration in reinforced concrete structures, a systematic in situ condition documentation, survey, and assessment of water tank structures has been done and presented in this paper, based on an empirical damage scale similar to that suggested in the other literature and a bilinear graphical deterioration model for such water retaining structures in a semitropical region like India on the basis of case studies. It contains in situ condition documentation, survey, and assessment of deteriorated structures which reflect the resultant deterioration process and also helps in validation of experimental and theoretical methods of performance evaluation. The water tanks of rural as well as urban areas are investigated considering correlation among structural type,

age, and condition rating, etc. Condition rating (CR) of water tanks is done on a scale of 0–9 prepared on the basis of similar scales proposed by Kaminetzky (1985) and Hollis and Gibson (1991). A deterioration scale for condition rating of water tanks has been designed with various classification levels.

The deterioration of structural component has been observed, where out of 204 tanks, 65% (132) are found to be badly deteriorated. Among these badly distressed tanks (132), 18% (24 of 132) are structurally deficient and functionally obsolete. Around 17% (22 of 132) of badly deteriorated water tanks are in serious condition. The average service life of water tanks, when first visible deterioration (e.g., cracks, etc.) were observed and maintenance was carried out, has been observed as 12 years and the average total service life with routine/minor repair and maintenance has been observed as around 46 years. However, in some cases the water tanks become structurally deficient in service life of around 25 years as shown in Fig. 5 & 6.

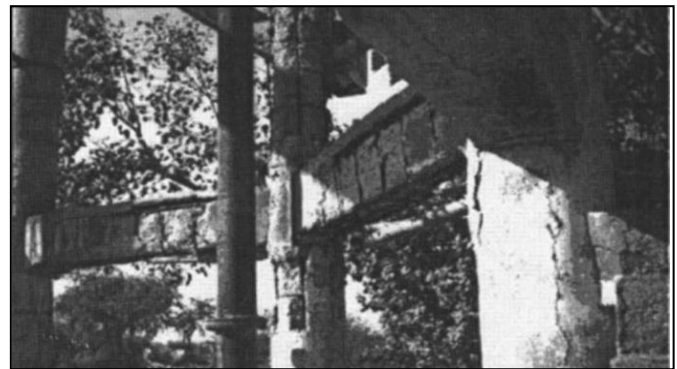


Fig. 5. View of deteriorated bracings and columns of water tank.



Fig. 6. Corrosion in reinforcement and spalling of concrete in bracings in water tank.

### D. Hemant and Sudhir (2007)

This paper reports on the effects of the great Sumatra earthquake and tsunami of December 26, 2004, in and around Port Blair, the capital city of the Andaman and Nicobar Islands, India caused widespread damage. It covers performance of overhead water tanks of Port Blair, where RC-framed staging are commonly used to distribute water to different parts of the city. During the shaking, some of these

tanks suffered substantial damages in the RC staging. A 50,000 L capacity water tank in the converter room of “Haddo naval jetty” is supported on 6-m high staging consisting of eight 300-mm square RC columns with intermediate RC beam braces at 3 m height (Fig. 7(a)). Plastic hinges developed at the bottom of all the columns and at the top of a few columns (Figs. 7(b) and 8, while the container sustained no damage. Reinforced concrete staircase was provided in the tank going around the staging, which seems to have made the tank geometry unsymmetrical, and the tank sustained a torsional response. An overhead water tank at MES headquarters, Birchgunj Military Station in Port Blair, was damaged by severe corrosion of reinforcement in the RC-framed staging and tank, but structural damage was not observed in the tank and it was still under use.

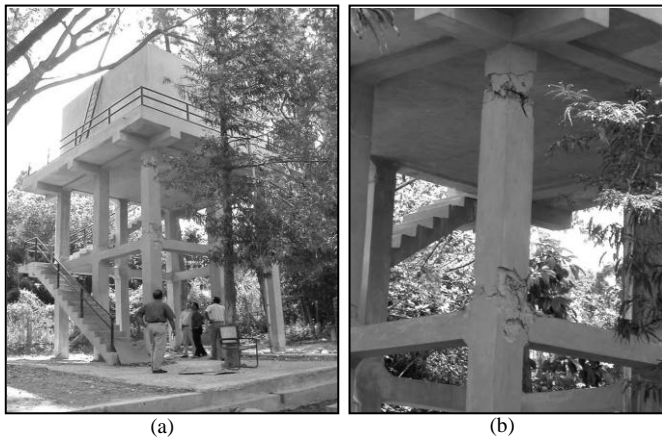


Fig. 7. (a) Water tank at Converter Room of Haddo naval jetty; (b) formation of plastic hinges at top of a few columns in tank staging

Fig. 8. Formation of plastic hinge at bottom of all RC columns in tank



staging.

#### E. Sudhir and Mahesh (2007)

In the present paper, a case study of deteriorated water tank structures situated in the semitropical region of India is presented. An investigation of failures of a series of water tanks of provincial states of Madhya Pradesh and Rajasthan in India, and the case study of anticipated amount of expenditure likely to be incurred in their repair and rehabilitation, further encouraged the research. Some selected parameters such as; concrete cover, carbonation depth, chloride concentration,

compressive strength, etc. which influence long term durability of structures have been measured.

The specific objectives of the present study are:

- To develop a methodology for field and laboratory performance testing and condition assessment of existing deteriorated water tank structures;
- To undertake a case study of field condition assessments of deteriorated water tank structures; and
- To make a comparative study of concrete cover, carbonation depth, compressive strength, and chloride concentration for different condition ratings (CR) based on field and laboratory performance testing.

A systematic study of water tanks of varying capacities between 50–2,500 KL constructed before 1995 around Gwalior, Bhind, Datia, Guna, Dholpur, etc. (districts of provincial states of Madhya Pradesh and Rajasthan of India) is conducted to establish significant correlations between structural type, age, condition rating, etc. Efforts have been made to trace the history of the tanks during the stages of construction and maintenance. The CR of water tanks has been done similar to scales proposed by Hollis and Gibson (1991), Kaminetzky (1985), and the Federal Highway Administration (FHWA). Field and laboratory measurements of compressive strength, concrete, cover, diameter of corroded reinforcement, carbonation depth, and chloride concentrations have been performed for water tanks of different CRs.

Due to corrosion, spalling of concrete had taken place which reduce the area of steel due to corrosion in bracing and columns, stairs, etc. observed in tanks under study is shown in Fig. 9 to 11, respectively. The failure of the water tanks has taken place mainly due to carbonation and chloride initiated corrosion because of the less concrete cover and measured carbonation depths exceeding clear concrete cover and chloride concentration exceeding safe corrosion threshold limits, as per Indian Standards, IS:456-2000 [11].



Fig. 9. Close-up of spalling of concrete in the bracing of one water tank





Fig. 10. Spalling of concrete and reduction of reinforcement cross section in bracing, column, and staircase of one water tank



Fig. 11. Vertical cracks and spalling of concrete in columns

*F. Amjad Masood, et al. (2008)*

The project launched by state government of Uttar Pradesh and implementation job given to the Corporation “Jal Nigam”, has been recorded with technology which was not appropriate for the Construction of the tanks and resulted in total failure of the project. Then, the project was referred to the Department of Civil Engineering, AMU, Aligarh for suggesting the remedy. The Department submitted a report highlighting the causes of extreme distress in the water tanks much earlier than their normal useful life of 50 years. A few tanks surveyed in rural areas of Agra and Mathura districts showed extreme damages in form of corrosion of reinforcement and spalling of concrete within a period of 10-15 years. It was observed that poor quality of concrete in terms of durability under the prevailing environmental conditions lead to the distress in water tanks. The deterioration of the water tanks due to distress is shown in (Fig. 12 & 13).



Fig. 12. Shaft at ground level near staircase

The project had several shortcomings in the planning phase, which lead to the above failures. It was further observed that the site selection of the project and location plays a very important role for the economy, safety and success of the project. The site selected was in the extreme interiors of the village where the human resources were rather scarce. Even the inflow of material to site was also not proper. The manpower hired was not trained properly for the mixing of concrete construction and this lead to poor workmanship and thus leading to poor quality of construction and thus poor structural performance of overhead water tanks.



Fig. 13. Shaft at ground level

The failure of the water tanks can thus be attributed to poor quality of concrete in terms of durability. The tanks surveyed for damage being in rural areas where it is difficult to maintain the quality of concrete. A durable concrete is one which does not show the signs of distress during its service life under given environmental conditions. Further the technology used for the construction was cast-insitu which in case of overhead water tanks is not suitable for remote areas, as this needs quality control of highest order. Lack of quality control has caused minor cracks in the water tanks, which lead to the leakage and deterioration of concrete. The reasons of failure thus are due to the improper selection of material, mix, placement, compaction, leaking formwork temperature control, curing leading to honeycombed and porous concrete. The honeycombing of concrete and improper cover to concrete, leads to the corrosion of reinforcement, because of the environmental attacks due to carbonation, chlorides and sulphate attack. The sulphate attack causes expansion and disruption of concrete due to formation of calcium sulphoaluminate. On the basis of the present study it has been concluded that, At project planning level varied human resources with diversified experience and knowledge should be engaged. Secondly, at planning and implementation level, local factors and psychology of people and their involvement in the project should not be ignored. Also, the factor of cost optimization at various stages of the project viz. designs,

procurement of material, project implementation, project startup and the subsequent operations should be done.

#### CONCLUSION

Generally, when earthquake occur major failures of elevated water tank take place due to failure of supporting systems, as they are to take care for seismic forces. Therefore supporting structures of elevated water tanks are extremely vulnerable under lateral forces due to an earthquake. A strength analysis of a few damaged shaft and frame type staging clearly shows them deficient, it shows current, IS: 1893-1984 underestimates the forces for water tanks and there are no suitable provisions delivered for ductile detailing of tank stagings though they are expected to undergo inelastic deformations during ultimate earthquake loads. So, staging requires such special attention towards selection of staging configuration, analysis and design to make elevated water tank structure live at the time of earthquake hazards. Nowadays, regular updating and revisions of codes are likewise necessary for improved design and construction practices in the country. As material, preventive measures can also be taken by using sulphate resisting cement or Ordinary Portland cement to avoid the sulphate attack.

Furthermore, it is required to take attention at project planning level for availability of varied human and equipment resources with diversified experience and knowledge. Also, life cycle costing, useful life and duration of the completion of the project is quite necessary to take in to consideration.

#### REFERENCES

- [1] Durgesh C. Rai, "Seismic Retrofitting of R/C Shaft Support of Elevated Tanks", *Earthquake Spectra*, vol. 18, no. 4, pp. 745-760, Nov. 2002.
- [2] Durgesh C. Rai, "Performance of elevated tanks in M<sub>w</sub> 7.7 Bhuj earthquake of January 26th, 2001", *Proc. Indian Acad. Sci. (Earth Planet Sci.)*, vol. 112, no. 3, pp. 421-429, Sept. 2003.
- [3] Sudhir S. Bhadauria and Mahesh C. Gupta, "In-Service Durability Performance of Water Tanks", *Journal of Performance of Constructed Facilities, ASCE*, vol. 20, no. 2, pp. 136-145, May 2006.
- [4] Hemant B. Kaushik and Sudhir K. Jain, "Impact of Great December 26, 2004 Sumatra Earthquake and Tsunami on Structures in Port Blair", *Journal of Performance of Constructed Facilities, ASCE*, vol. 21, no. 2, pp. 128-142, March/April 2007.
- [5] Sudhir S. Bhadauria and Dr. Mahesh C. Gupta, "In Situ Performance Testing of Deterioration Water Tanks for Durability Assessment", *Journal of Performance of Constructed Facilities, ASCE*, vol. 21, no. 3, pp. 234-239, May/June 2006.
- [6] V. P. Mital, "Failure of Overhead Water Tank in the State of Uttar Pradesh in India – A Case Study", *First Int. Conf. Construction in Developing Countries (ICCIDC-I)*, pp. 185-191, August – 2008.
- [7] IS: 1893-1984 (part-2), "Criteria for Earthquake Resistant Design of Structures", Bureau of Indian Standards, New Delhi.
- [8] IS: 11682-1985, "Criteria for Design of RCC Staging for Overhead Water Tanks", Bureau of Indian Standards, New Delhi.
- [9] IS: 13920-1993, "Code of Practice for Ductile Detailing of Reinforced Concrete Structures subjected to Seismic Forces", Bureau of Indian Standards, New Delhi.
- [10] IS: 4326-1976, "Code of Practice for Earthquake Resistant Design and Construction of Buildings", Bureau of Indian Standards, New Delhi.
- [11] IS: 456-2000, "Code for Practice for Plain and Reinforced", Bureau of Indian Standards, New Delhi.

# Multilevel Car Park – Optimum Solution While Considering Composite Option

A. Chirag N. Patel, B. Savan B. Kakadia

**Abstract**— Paper presents multilevel car park as the right solution to meet the increased demand of parking facility. The car parks made of composite construction are ideal solution since they are economical and offer large column free area apart from lot of advantages due to its lightweight. It has got more earthquake resistant properties and large column free areas for free movement of traffic. Various aspects of multilevel car parks and general requirements of multilevel car parks with its structural and functional arrangements are discussed. As a functional point of view, the split level type ramp system is used because this system is more efficient for user to drive and park the car in parking stall. The steel decking floor system has been designed using BS 5950 (Part-4). Analysis of moment resisting and braced frame has been compared using STAADpro, and designed with working stress method. The results of deflection, bending moment, weight difference have been studied. The study shows that the multilevel car park construction is cost effective while considering composite construction. The result of analysis shows that the saving in weight while the braced frame is used rather than the moment resisting frame. Also, the profile decking floor slab is time effective solution. Hence, the composite construction can be adopted for saving in time and in cost also.

**Index Terms**—Multi level, Car Park, Composite.

## I. INTRODUCTION

Nowadays vehicular traffic in the metropolitan cities has been expanding at a very fast rate. It is now poised for greater growth as the country's economy enters take off stage. Many new companies have started manufacturing cars in India to cater market of Indian society. According to today's scenario, more and more people can afford to buy cars. This upsurge in vehicles has created a big problem of parking particularly in congested commercial and office localities therefore concept of multilevel car parks has become a need of the day. The multi-storey car park is an exceptional style of building, one in which all elements of the structure are normally exposed to the environment. One must remember that these car parks must be completed quickly and without causing much hindrance to the busy traffic.

As early as 1918, pioneering Chicago innovate new architectural designs to keep cars. Holabird and Roche

A Assistant Professor of Applied Mechanics Department, L. D. College of Engineering, Ahmedabad, Gujarat, India. (e-mail: cnpatel.693@gmail.com)

B

Consulting Structural Engineer of Bhagavati Engineers, Rajkot, Gujarat, India. (e-mail: kakadiasavan@gmail.com)

designed a vertically stacked 5-story structure with a spiraling ramp for the Hotel "La Salle".

## II. TYPES AND ARRANGEMENTS OF CAR PARKING STRUCTURE

### A. Operational Types

- Automated park: It requires half the volume of a conventional car park system. Due to Steel-framed car parks, it doesn't require access ramps or roadways within the car storage area. The driver parks the cars on a robot trolley within an entrance module. From this point the trolley takes the car to an empty parking space.
- Self-park: In this facility, driver drives the car from street to the stalls without any obstruction. For exit and entrance the different stair or lift arrangement provided for the driver.

### B. Material Types

- Reinforced Concrete Construction
- Steel-Concrete Composite Construction
- Steel Construction

### C. Functional Types

- Twin-spiral Type: The ramps, situated in opposite corners, are angled to facilitate the movement from the floors to the ramps. Entrances and exits have been provided on separate levels to take advantage of the different elevations.
- Split-level type or staggered floor type: The ramp systems feature separated one-way operation, and access is on only one street. Ninety-degree parking is utilized throughout the four floors Fig.1.

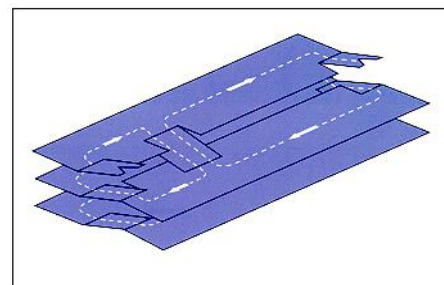


Fig.1 Split-level type

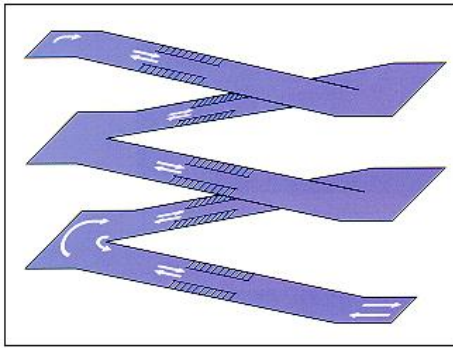


Fig.2 Sloping floor

- Straight ramp type: Straight ramp is provided. A portion of the aisles is used in the floor-to-floor circulation. The widths of the ramps should not be less than 3.65 m for a single ramp and 7.0 m for a double ramp.
- Spiral ramp type: The preparation of a design for an irregular shape site presents many problems, especially when self-parking is to be provided. At that time this type of system can be used.
- Sloping floor or continuous-ramp type: The aisles serve two purposes: access to the parking stalls, and floor-to-floor circulation Fig.2.

The main objective of present work is optimum solution using composite construction. Comparison of RCC and steel-composite construction studied in INSDAG publication [5], as per that the steel-composite construction gives optimum solution over the RCC construction above the three storey parking and also it gives large column free area. So here, attempt has been made to achieve optimum solution for multi-level car park, while adopting composite construction and making comparison of different types of structural system.

### III. CONSIDERATION OF FRAME SYSTEM

In the present study Split-level type or Staggered floor system have been considered with Moment Resisting Frame (MRF) and Centrally Braced Frame (BF) Systems.

#### A. Moment Resisting Frame (MRF):

A moment frame structure consists of columns and girders jointed by moment-resistant connections. The lateral stiffness of a moment frame bent depends on the bending stiffness of the columns, girders and connections in the plane of the bent. The principal advantage of moment frame is its open rectangular arrangement, which allows freedom of planning and easy fitting of doors and windows.

Gravity loading also is resisted by the moment frame action. Negative moments are induced in the girders adjacent to the columns causing the mid-span positive moments to be significantly less than in a simply supported span.

#### B. Centrally Braced Frame (BF):

The braced frame is a common system employed to resist the significant lateral loads that exceptionally tall structures are subjected to. The advantages of braced frames from a structural engineering standpoint are enormous. Braced frames carry the lateral forces in an axial manner than through the bending of elements which is highly inefficient. The largest drawback for braced frames is that the scheme is obstructive and significantly reduces openings within bays. But in the case of multilevel car parks, the requirement of opening is not much. However, it should be noted that while the bracing element covers a significant portion in criteria bays, the separation of the lateral system from vertical system leads to much large column spacing which allows more flexibility in programming of the interior space within those bays.

### IV. GENERAL CONSIDERATION FOR MULTILEVEL CAR PARKS

Based on popularity of car parks in U.K. five storied split level type car park considered for achieving optimum solution using composite option.

#### A. Design Parameters:

The design parameters considered for design are as follows:

- Built-up area : 80 m x 32 m
- No. of floors : 5 (Including Ground/basement)
- Parking space : 2.5 m x 5 m (per car)
- Width of ramp : 5 m (2 stall widths)
- Inclination of ramp : 14% (For split level type)
- Height between floors : 3.1 m (All floors)
- Clearance height : 2.33 m
- Safe soil bearing capacity of soil : 300 kN/m<sup>2</sup>
- Importance factor : 1
- Earthquake zone : III
- Response reduction factor : 3
- Slab thickness : 130 mm
- Wind speed : 44 m/s

#### B. Design Codes and Methodology:

The design reference codes considered for design are as follows:

- BS: 5950 (Part 3.1 & Part 4) [8]
- Euro Code 3 (Part 1.1) [12]
- Euro Code 4 (Part 1.1) [13]
- IS: 800-1984 [11]
- IS: 1893-2002 (Part 1) [6]
- IS: 875-1987 (Part I & II) [9], [10]
- IS: 11384-1985 [4]
- IS: 456-2000 [7]

The design methodology used for structure are as follows:

- Floor is prepared using profile steel deck, with topping of floor finish.
- All beams, columns and bracings are fabricated from steel.
- Profile steel deck and beams has been designed by the limit state method using partial safety factor for loads

and material strength specified in BS: 5950 (Part 4) and EC4 respectively.

- Shear Connector: Stud shear connector has been used and capacity is taken from IS: 11384-1985
- Composite beam design is made as per BS: 5950 (Part 3.1).
- Cambering is required in all 16m span beams to take care of deflection serviceability.
- Facilities provided : Attendant’s room and toilet at ground floor, Two nos. lifts, Two stair case, Parking for about 558 no. of cars

V. GENERAL REQUIREMENT OF SPLIT LEVEL TYPE CAR PARKS

The ramp systems feature separated one-way operation, and access is on only one street. Ninety-degree parking is utilized throughout the floor. The cashier’s booth is at the entrance and the stairs and elevators are strategically located to take advantage of the split-level and to afford minimum walking distances.

TABLE.1 INTER FLOOR DISTANCE

Slope	Floor to Floor Height					Half Height Split Level	
	8 ft	9 ft	10 ft	11 ft	12 ft	4 ft	5 ft
5%	160	180	200	220	240	80	100
6%	114	128	153	157	172	57	77
7%	67	75	83	92	100	34	42
8%	57	64	72	79	86	29	36

Split-level floor construction requires the length of ramp travel to be about one-half the usual inter-floor distance shown in Table 1. This was one of the most common designs for years. Split-level floors can overlap as much as five to six feet at the split, which increases space efficiency and makes a narrow site workable.

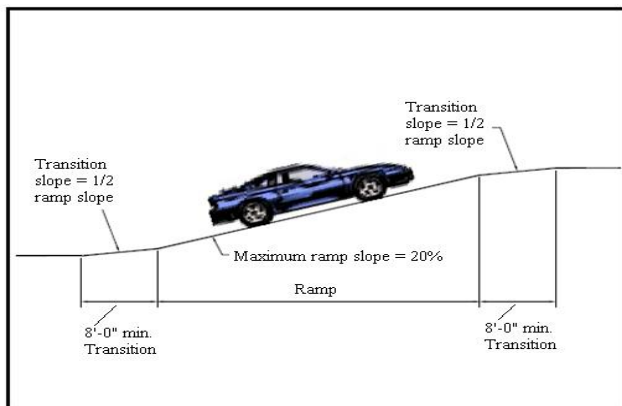


Fig.3 Ramp slope and transition

Typically, two ramps are utilized supporting either one- or two-way traffic. Parking and walking does not take place on the ramps. Ramps can be sloped as greatly as 16 percent, but any slopes over 14 percent will require transitions seen in Fig.3. Each ramp climbs half the height of the tier.

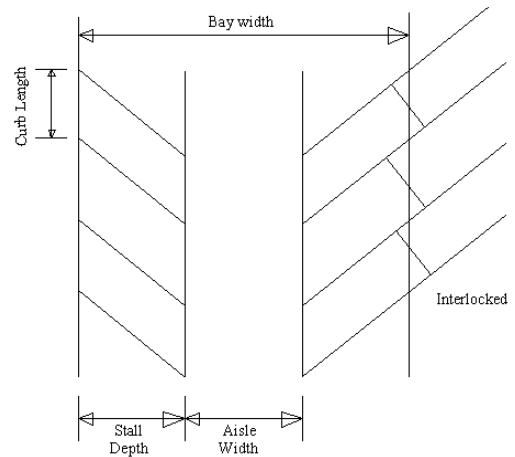


Fig.4 Interlocked system parking

TABLE.2 PARKING ANGLE AND DIFFERENT SIZES

Elements	Park Angle (Inter Locked)			
	45°	60°	70°	90°
Curb Length (ft)	12.33	10.10	9.33	8.75
Depth of Stall(ft)	13.60	16.60	17.50	17.75
Aisle Width(ft)	13.67	15.50	17.50	25.00
Bay Width(ft)	40.87	48.70	52.50	60.50

For utilizing maximum space on the floor, the interlocked system can be used. In the interlock system the car parked at angle and it interlocked the stall depth seen in Fig. 4 & Table 2.

Every parking structure is required to have a minimum of two means of egress (stairs) which are separated from each other. These stairwells should be located based upon the requirements of local safety codes. One of the stairwells is normally located adjacent to the elevator locations. If the parking structure supports a high peak flow, extra wide stairs may be necessary. The minimum clear width for the stairway is 36 in. (0.9144 m). Stairwells on the perimeter of the structure are often left open or glass enclosed to create a sense of security within the structure.

The number and locations of stairs is determined by the maximum travel distance or the distance the patron must travel along a normal path from any point in the structure to the closest stair. In open structures the controlling distance is 300 ft. (91.44 m).

The number of elevators required to service the parking structure varies by the usage of the structure. Elevators should be located along the natural direction of travel for a patron exiting the structure. Some parking designers take into account elevator locations in specifying the traffic flow through the facility. In doing so, all incoming traffic is routed past the elevators as a means of orienting the patron to the structure layout.

VI. GEOMETRY AND LOAD CONSIDERATION FOR MRF AND BE SYSTEMS

From the literature survey, the plan of multilevel car parking is prepared and selects the split-level ramp system. In this system the four single ramps are provided. Two ramps are place at extreme end of rectangular long side and two are placed at the intermediate of previous two. Also the two staircase and two elevators are placed at exterior side of the end ramps. The cashier’s room and toilet unit are provided at left and right side of entrance.

The design of floor system is carried out using the profile steel decking and also the economic as well as time dependent aspects have been discussed. The composite beam is designed considering profile steel decking above it. Also the design of column has been done using STAADpro software. The design results have given of moment resisting frame system in tabular form of each and every floor.

The following figures indicate the general arrangement, plan, elevation and cross section details including marking of beams, columns and bracing of 5 levels split level car park considered for design.

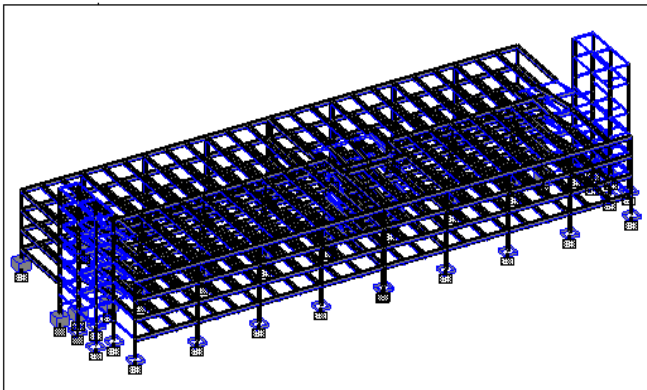


Fig.5 Sectional 3D view of Multilevel Car Park (MRF system)

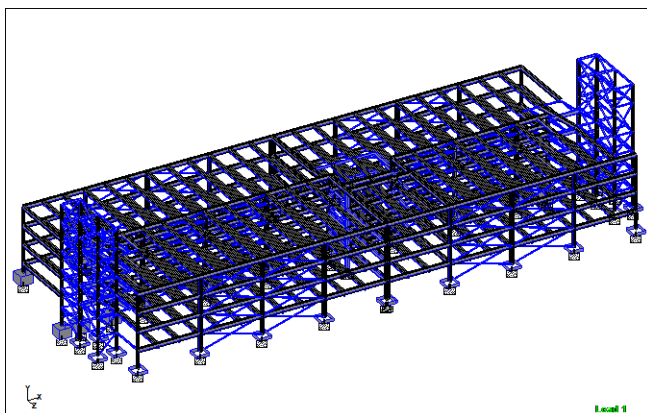


Fig.6 Sectional 3D view of Multilevel Car Park (BF system)

Analysis has been carried out using STAADpro software. 3D geometrical view of multilevel car park structure which is taken for analysis is shown in Fig. 5 and 6 for MRF and BF respectively.

A. Load Consideration

- Earthquake Load: As per IS: 1893-2002[6], the seismic coefficient method has been used the evaluation of the earthquake forces.
  - Live Load (From Vehicular wheel load) :
  - Uniformly Distributed Load 2.5 kN/m<sup>2</sup>
  - Concentrated Load at any point 9.0 kN
    - Lift Load 10 kN/m<sup>2</sup>
    - Staircase load, taken as UDL on adjacent beam
    - Wall Load 20 kN/m<sup>3</sup>
- (Wall provided around the lift block)

VII. COMPARISON OF FLOOR SYSTEM

Study says that above three storeys, the composite construction is economical for multilevel car parks. Literature survey shows that in most multilevel car parks the steel framing is used. Purpose behind the use of steel framing is saving the time and achieving the fast construction. So here, the comparison has been carried out considering the two framing system. One is moment resisting frame system and other is braced frame system. For the braced frame, it is find that the x-bracing is economical over the other types of bracing system. And it is better solution for resisting a lateral loads which coming from the either direction.

Based on study, the composite floor using profile sheeting has been considered as the better solution for the floor construction of multilevel car parks. If we considered only the cost than it is more than the simple solid RCC slab but in consideration of time saving it is highly economical.

A G+5 commercial building considered for study is a rectangular (48 x 13.5 m) in plan with nominal height of 16.2 m (2.7 m floor to floor) and gross floor area of 3900 m<sup>2</sup> (650 m<sup>2</sup> area at each floor) constructed at pune. For considering same total floor area, height and loading condition, this building is designed and constructed by two different methods. It is found that the time saving of 55.29% is achieved due to use of composite floor construction rather than RCC floor (Table 3). The direct construction cost required for composite floor is 27.64 % higher than RCC floor. But overall the net cost required for composite floor is only 0.45 % more than RCC floor considering time related saving (Table 4 & 5).

TABLE 3  
FLOOR TO FLOOR TIME SAVINGS

Slab Level	Duration in (Days)		Time Saving (Days)	Time Savings (%)
	Composite Floor	RCC Floor		
1	7	15	8	53.33
2	8	17	9	52.94
3	9	19	10	52.63
4	10	21	11	52.38
5	10	24	14	58.33
6	11	27	16	59.26

Total	55	123	68	55.29
-------	----	-----	----	-------

TABLE 4  
FLOOR TO FLOOR TIME COMPARISON BETWEEN RCC & COMPOSITE FLOOR CONSTRUCTION

Slab Level	RCC Floor Construction				Composite Floor Construction						
	Erection of Slab & Beam Formwork (Days)	Lifting & Laying of Steel Reinforcement (Days)	Concreting (Days)	Total Duration (Days)	Lifting of Steel Decks (Days)	Placing & Installation of Steel Deck (days)	Cleaning of Steel Decks (Days)	Lifting & Placing of Reinforcement (Days)	Concreting (Days)	Total Duration (Days)	Erection of Slab & Beam Formwork (Days)
1	8	6	2	15	1	3	1	1	1	7	1
2	9	7	2	17	2	3	1	1	1	8	2
3	10	8	2	19	2	3	1	2	1	9	3
4	11	9	2	21	3	3	1	2	1	10	4
5	13	10	2	24	3	3	1	2	1	10	5
6	15	11	2	27	4	3	1	2	1	11	6
Total duration for RCC floor construction = 123 days					Total duration for Composite floor construction = 55 days						

TABLE 5  
NET COSTS FOR RCC & COMPOSITE FLOOR CONSTRUCTION

Cost	RCC Floor Construction	Composite Floor Construction
Direct Cost	6846150	8737210
11% interest on direct cost (For RCC: 123 days & For Composite: 55 days)	253780	144825
10% interest on 1 <sup>st</sup> installment of selling amount (for 68 days)	1077686.3	--
10% interest on 2 <sup>nd</sup> installment of selling amount (for 38 days)	451677.3	--
10% interest on 3 <sup>rd</sup> installment of selling amount (for 18 days)	213952.4	--
Total net Cost (Rs)	8843246	8882035
Total net Cost (Rs/m <sup>2</sup> )	2267.5	2277.5
Extra Cost required for Composite over RCC floor	10	

considering net cost (Rs/m <sup>2</sup> )	
Extra Cost required for Composite over RCC floor considering net cost (%)	0.45(39000)

TABLE 6  
DEFLECTION COMPARISON

SL	Description	Moment Frame		Braced Frame	
		X	Z	X	Z
1	Top of Lift Room	60 mm	42 mm	36 mm	24 mm
2	Top Floor	52 mm	18 mm	21 mm	21 mm

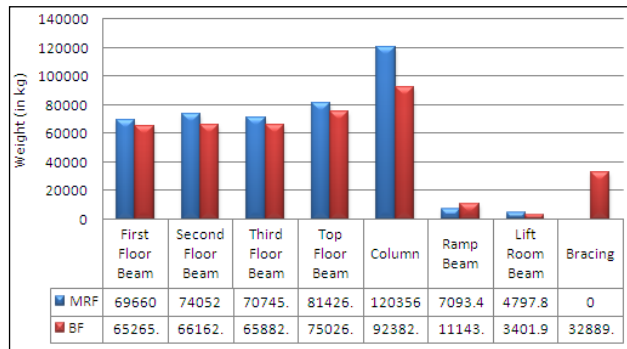


Fig. 7.1 Weight comparison

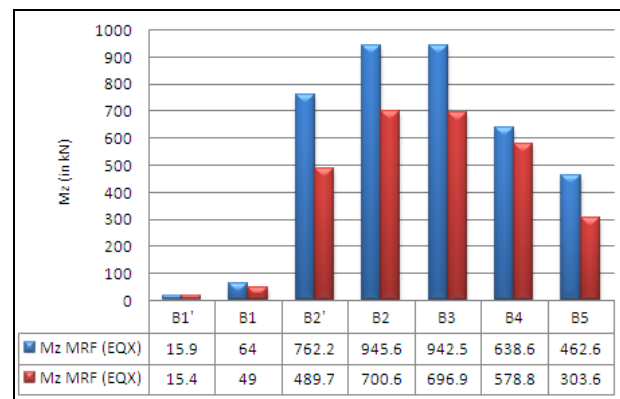


Fig. 7.2 Foundation moment comparison

VIII. SUMMARY

Multilevel car parks have become necessity in many metropolitan cities, due to insufficiency of land. Hence multilevel car park is the right solution to meet the increased demand of parking facility. The car parks made of composite construction are ideal solution since they are economical and offer large column free area apart from lot of advantages due to its lightweight. It has got more earthquake resistant properties and large column free areas for free movement of traffic. The report contains the various aspects of multilevel car parks. It gives general requirement of multilevel car parks with its structural and functional arrangement.

As a functional point of view, the split level type ramp system is used because this system is more efficient for user to drive and park the car in parking stall. The floor system considered of profile steel decking. The design of decking floor system has been carried out using BS 5950 (Part-4). The seismic analysis has been carried out using IS 1893 (Part-2). The steel framing is used in the study while the dead load and live load are considered from IS 875 (Part-2). The comparison of two structural systems has been done like one system taken as moment resisting frame and other is braced frame. Analysis of above two systems has been done using *STAADpro 2006*. Design has been carried out using working stress method. The results of deflection, bending moment, weight difference have been studied.

### CONCLUSIONS

The study shows that the multilevel car park construction is cost effective while considering composite construction. The result of analysis shows that the saving in weight while the braced frame is used rather than the moment resisting frame. Also, the profile decking floor slab is time effective solution. Hence, the composite construction can be adopted for saving in time and in cost also. There are some other conclusion is:

- The 3.7% weight reduces while considering braced frame over the moment resisting frame.
- The reduction in the deflection is also observed.
- The load carrying behavior (Through element bending) of moment frame results in significant column and girder end moment.
- This larger moment leads to design the larger section
- Due to higher design end moment the connection cost will be affected
- The bending moment is reduced in column and beam when bracings are provided.
- Due to reduction in bending moment, the smaller section can be provided. Also it reduces connection cost.
- The bending moment gets reduced in foundation also. So it gives lighter foundation while bracing is used.
- The construction of profile deck floor system is saving a half the time over the solid concrete slab.
- Also the profile deck floor system is lighter than the solid concrete slab system and this reduction in weight will affect the foundation cost.

### REFERENCES

- [1] Pydi Rao, "Comparative study of multilevel car parks with RCC and composite option".
- [2] Roy Becker, & Michael Ishler, "Seismic design practice for eccentrically braced frames", Steel Tips, December, 1996.
- [3] IS 11384:1985, "Code of practice for composite construction in structural steel and concrete" Bureau of Indian Standards, New Delhi, 1985.
- [4] Institute for Steel Development & Growth, "Hand Book on Composite construction (Multilevel car parking)", and INSDAG publication no ins/pub/019, January, 2002.
- [5] IS 1893 (Part-1):2002, "Criteria for earthquake resistant design of structures" Bureau of Indian Standards, New Delhi, 2002.
- [6] IS 456:2000, "Plain and reinforced concrete- Code of practice" Bureau of Indian Standards, New Delhi, 1995.
- [7] BS 5950 (Part-4):1994, "Structural use of steel work in building, code of practice for design of composite slabs with profiled steel sheeting" British Standard.
- [8] IS 875 (Part-1):1987, "Code of practice for design load for building and structure- Dead Load" Bureau of Indian Standards, New Delhi, 2002.
- [9] IS 875 (Part-2):1987, "Code of practice for design load for building and structure – Imposed Loads" Bureau of Indian Standards, New Delhi, 1989.
- [10] IS 800:1984, "Code of practice for general construction in steel" Bureau of Indian Standards, New Delhi.
- [11] Eurocode-3, "Structural steelwork euro code", composite structure guideline given, European Standard.
- [12] Eurocode-4, "Design of composite steel and concrete structures - general rules and rules for building", European Standard.



# A Combined Subspace Method for Face Image Retrieval

A. Fousiya K. K, B. Jahfar Ali P.

**Abstract-** Face recognition and face image retrieval systems have attained much importance in recent times, due to many real time application demands it. Normally, image retrieval has become a challenging issue in the real world applications because the vector space representation of image collections are often in high dimensional. This high dimensionality causes a degradation in the efficiency of a image retrieval system. The dimensionality reduction has a significant impact over the performance classifications algorithms based on machine learning and data mining techniques, which plays an vital role in image retrieval systems. With regard to the machine learning approach, image retrieval can be viewed as the problem of classification. By providing better dimensionality reduction and higher class discrimination prior to the classification process leads to higher classification accuracy, which ultimately creates better retrieval rate. In this work, a face image retrieval system which retrieves the relevant face images from a large face database with response to a query face image is developed. Face recognition is the most important process of the face retrieval systems. The major step of face recognition is feature extraction and classification. The performance of a face recognition system highly depends on the way in which the features are extracted and the classification of them to the appropriate group. This is achieved by merging Principal Component Analysis, that provides better dimensionality reduction and Linear Discriminant Analysis, that provides better class discrimination which is advantageous for improving classification accuracy of Support Vector Machine classifier. The experiments are implemented on the Olivetti Research Laboratory (ORL) face database, which provides higher variability in facial expressions, pose and facial details. This method improved the classification accuracy.

**Index Terms**—Dimensionality reduction, Face recognition, PCA, LDA, SVM.

## I. INTRODUCTION

Face image retrieval system retrieves the face images from a large face database, which are similar to a query face image. Studies shown that, from a machine learning perspective image retrieval can be viewed as a problem of classifying images into the appropriate classes [10]. Therefore, face image retrieval directly depends on the classification the

A. M. Tech Scholar, MES College of Engineering, Kuttippuram, Kerala, (e-mail: fousiadath@gmail.com).

B. Asst. Professor, MES College of Engineering, Kuttippuram, Kerala, (e-mail: jahfar.ali@gmail.com).

face images to the proper classes and so, face recognition is the major step in the face retrieval system. So, the retrieval rate directly depends on the classification accuracy. Face recognition system is a major issue in applications such as security systems, surveillance and credit card verification. Face recognition is the process of recognizing human faces with many variations in facial appearances such as facial expression, poses and facial details.

Normally, the face images are high dimensional. So, face recognition has to address the dimensionality reduction problem because, a face image having  $m*n$  pixels is represented by a vector in  $R^{m*n}$  space for computational purposes. This  $m*n$  dimensional space is too large for recognition process. When the number of images in the data set increases, the complexity of representing the data set also increases. Analysing the data set with a large number of variables increases the computational complexity and consumes a large amount of memory. When dealing with machine learning and data mining algorithms, this high dimensionality causes a degradation in the efficiency of a classification system. Thus, for solving the curse of dimensionality, various dimensionality reduction methods have been applied as a preprocessing step.

Often, the images including digital photographs, scientific images, medical images, hyper spectral images and graphic images are high dimensional and there exist hundreds or even thousands of features. When use with machine learning and data mining algorithms, these larger number of features lead to major problem known as the curse of dimensionality. Curse of dimensionality causes degradation in the efficiency and performance of the system. For the efficient processing of high dimensional images, its dimensionality has to be reduced without a loss in the original properties of the high dimensional space. Thus, for solving the curse of dimensionality, various dimensionality reduction methods have been introduced.

The remainder of this paper is organized as follows. In section II, describe about dimensionality reduction techniques. Section III presents related works. Section IV describes about face retrieval system. Section V goes through the experimental results and conclusion is described in section VI.

## II. DIMENSIONALITY REDUCTION METHOD

Dimensionality reduction is the transformation of high dimensional patterns into a lower dimensional representation. The goal of the dimensionality reduction is to preserve the

local structure of the original high dimensional space. Dimensionality reduction removes the irrelevant and redundant features, and extracts a small number of relevant features. The minimum number of parameters which are required for representing the properties of the data is known as the intrinsic dimensionality. So, the dimensionality of the reduced representation should corresponds to the intrinsic dimensionality. The lower dimensional representation provides the efficient visualization of the data and analysis of the high dimensional data can be done efficiently by projecting them into a lower dimensional representation. By dimensionality reduction, irrelevant features of high dimensional data are eliminated. For example, the images produced by the hyper spectral satellites contain hundreds of spectral bands. For remote sensing applications, these higher bands are a major problem and the preprocessing requires a reduction in the number of bands, which is shown in Figure 1.1 [4].

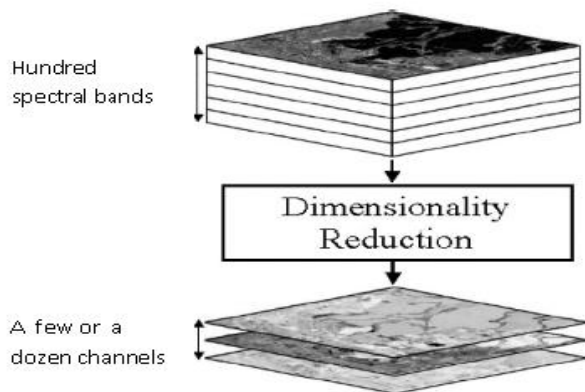


Figure 1: Preprocessed dimensionality reduction prior to information derivation from hyper spectral imagery.

Dimensionality reduction techniques are classified into linear and non-linear techniques. A linear method transforms the high dimensional data into a lower dimensional subspace using a linear mapping. Linear dimensionality reduction techniques finds a lower  $k$  dimensional representation  $S$  for the original  $p$  dimensional data  $X$ , where  $k < p$  and each variable of the reduced representation is a linear combination of the original variables as given in (1) or in (2).

$$S_i = w_{i,1} x_1 + \dots + w_{i,p} x_p, \text{ for } i=1,2,\dots,k \quad (1)$$

$$S = W_X \quad (2)$$

where  $W_{k \times p}$  is the linear transformation weight matrix [2]. Among the previously investigated methods for linear dimensionality reduction, the most popular are: Principal Component Analysis (PCA) [4] and Linear Discriminant Analysis (LDA). The linear techniques are simple and easy to implement.

The non-linear methods like Locally Linear Embedding (LLE) [5], Isomap [6] handle the complex non-linear structures efficiently. In some cases, the observed data lies on an abstract space called manifold of the original high dimensional space. For example, the variations that occur in the face images can be represented using a low dimensional face manifold, lying on a high dimensional image space. This

form of non-linear dimensionality reduction methods are known as manifold learning techniques [2][3]. Manifold learning extracts the non-linear structures or low-dimensional manifolds embedded in the high dimensional data sets and this low dimensional embedding preserves the structures that exist in the original space. In such cases, the non-linear dimensionality reduction techniques can find the manifold structure efficiently, whereas the linear techniques fail to do so.

### III. RELATED WORK

Recently a significant number of works have been done on face detection and recognition. Face recognition methods can be classified as feature based methods and subspace methods. Feature-based approaches extract the local features such as the position of the eyes, nose, mouth etc. Subspace method reduces the dimension of the data, while retaining the maximum separation between distinct classes.

'Eigenface' [13] and 'Fisherface' [4] are the two widely used subspace methods for face recognition. The 'Eigenface' approach uses the linear unsupervised dimensionality reduction method Principal Component Analysis (PCA) for subspace generation, whereas the 'Fisherface' approach uses linear supervised dimensionality reduction method Linear Discriminant Analysis (LDA).

PCA finds a projection on a lower dimensional representation, where along the principal component most of the data variation occurs in an unsupervised manner. PCA provides the accurate representation of the data with minimum reconstruction error and also finds the best axis for projection [4]. The main aim of LDA is to maximize the discrimination between different classes, while minimizing the within class distance. In classification systems, LDA is superior to PCA because, it provides higher class discrimination by using the class information [3][4] and so, LDA is widely used in face recognition systems. Class discrimination capability of PCA is very less compared to LDA [3].

Mahesh Pal and Giles M. Foody [7] discussed about the need of dimensionality reduction as a preprocessing step to classification. Support Vector Machine (SVM) is a widely used method for classification. SVM constructs a hyper plane in which the margin between different classes should be high and so, it provides better classification accuracy. Experiments are conducted for classification of hyper spectral data using SVM. Experimental results show that the classification accuracy of SVM can be increased by reducing the dimensionality of the data [7]. Thus, the authors proved that, dimensionality reduction is an essential pre-processing stage for classification by SVM. Besides providing classification accuracy, dimensionality reduction offers the following advantages to SVM:

- Improved speed of the classification by reducing feature set size
- Reduction in data storage requirements

Jianke Li et al. [8] introduced a method for face recognition system based on the combination of Principal Component Analysis (PCA) and Linear Discriminant Analysis

(LDA) for feature extraction. In order to perform feature extraction, a commonly used method is PCA and it provides better dimensionality reduction. After performing the PCA transformation, the resulting components are uncorrelated and have less reconstruction error. LDA increases the class separation and extracts the LDA features, which is advantageous for classification. So, for providing higher discrimination of the samples, LDA is used after PCA transformation. Thus PCA forms a PCA subspace and the combination of PCA and LDA form the LDA subspace having high discrimination. For classification purpose, Nearest Neighbor Classifier (NNC) is used. Experiments are conducted on the ORL face database containing 400 images of 40 individuals. Classification accuracy for different training sample size and for various feature dimensions was calculated. An increase in training sample size causes an increase in the classification accuracy of both PCA method and combined PCA and LDA method. Also, it can be concluded that the PCA and LDA combination method provides better classification accuracy than the PCA alone method for face recognition [8].

In [9], Md. Omar Faruque et al. proposed a face recognition system using Principal Component Analysis (PCA) and support vector machine (SVM). Since the major components of a face recognition system are feature extraction and classification, its performance strongly depends on the factors including the way in which the features are extracted, how accurately the features are classified into a group and the selection of a classifier. Therefore, the selection of the feature extractor and the classifier is very important in a face recognition system. In this work, the authors have used PCA for feature extraction and SVM for classification.

A classification system provides higher classification accuracy only if it is preceded by a better dimensionality reduction stage and better class discrimination stage. From the literature survey, it is observed that whatever be the classifier used for classification, dimensionality reduction is an essential step prior to the classification stage [7]. Table I. summarizes the different dimensionality reduction methods used for face recognition in the literature survey. Experiments are conducted on the ORL database containing 400 images of 40 individuals.

TABLE I. CLASSIFICATION ACCURACY VS FEATURE DIMENSION

Method	Classifier	Feature Dimension	Classification Accuracy
PCA	NNC	40	85%
PCA + LDA	NNC	40	93%
PCA	SVM	40	97%

#### IV. FACE RETRIEVAL SYSTEM

The proposed face retrieval system has two major components: face recognition system and retrieval system.

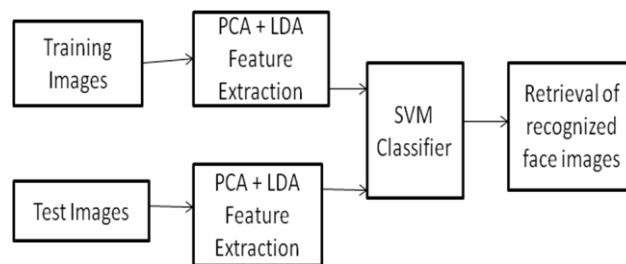


Figure 2: Proposed face retrieval system model

A face recognition system consists of two major steps as shown in Fig. 2.

- (1) Feature extraction
- (2) Classification

Feature extraction is process of reducing the high dimensional training data into a set of features to understand about the structure of the data. The feature extraction step extracts the features of the face images. Then the classifier is trained for training images. Now the extracted features are classified using a classifier and so, the classifier classifies the face images to the appropriate classes.

The first step of human face recognition is the extraction of the relevant features from facial images. Then the extracted features are classified using a classifier. Here, the combination of PCA and LDA is used for feature extraction and redundancy eliminated SVM for classification. Figure 3 represents the model of the face recognition system.

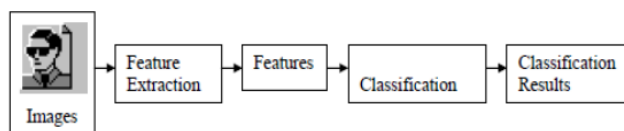


Figure 3: A face recognition system model

##### A. Feature Extraction Using PCA

PCA reduces the dimension of the data by finding a few orthogonal linear combinations known as principal components of the original variables with the largest variance. The first principal component is the linear combination with the largest variance. The second principal component is the linear combination with the second largest variance and it is orthogonal to the first principal component and so on. The number of principal components is same as the number of the original variables. For many datasets, the first several principal components explain most of the variance, so that the rest of the principal components can be discarded with minimal loss of

information. Sample data is approximately reconstructed by using only part of their projection onto the PCA subspace. The steps of face recognition using PCA are as follows:

- Step 1: Obtain face images  $I_1, I_2, \dots, I_M$
- Step 2: Represent every image  $I_i$  as a vector  $\Gamma_i$
- Step 3: Compute the average face vector  $\Psi$ :

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (3)$$

Step 4: Subtract the mean face:

$$\Phi_i = \Gamma_i - \Psi \quad (4)$$

Step 5: Compute the covariance matrix C:

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T \quad (5)$$

Where  $A = [\phi_1, \phi_2, \dots, \phi_M]$

- Step 6: Compute the eigen vectors  $\lambda_i$  of  $AA^T$
- Step 7: Keep only K eigen vectors. Choose the following criterion for the selection of K.

$$\frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^N \lambda_i} > \text{threshold value} \quad (6)$$

### B. Feature Extraction Using LDA

Basic step of LDA algorithm is as follows:

Step 1: Determine LDA subspace from training data. For that find within class scatter matrix  $S_W$  and between class scatter matrix  $S_B$ .

$$S_W = \sum_{j=1}^C \sum_{i=1}^{N_j} (x_i^j - \mu_j)(x_i^j - \mu_j)^T \quad (7)$$

Where  $X_i^j$  is  $i$ th sample of class  $j$ ,  $\mu_j$  is the mean of class,  $C$  is the number of classes,  $N_j$  is the number of samples.

$$S_B = \sum_{j=1}^C (\mu_j - \mu)(\mu_j - \mu)^T \quad (8)$$

Where  $\mu$  is the mean of all classes.

Step 2: Project all training images onto the LDA subspace.

Step 3: Project each test image to the same subspace and compare by using the distance metrics between the image and training images.

LDA searches for the directions that provide maximum discrimination of classes in addition to dimensionality reduction. This is achieved using within-class and between-class matrices. Within-class scatter matrix is the scatter of the samples around their respective class means and between-class scatter matrix is the scatter of class means around the mixture mean. The aim of LDA is to maximize between class data separation and minimize within class data separation. Various

measures are available for quantifying the discriminating power. Commonly used one is Fischer criterion as:

$$J(W) = \frac{|W^T S_B W|}{|W^T S_W W|} \quad (9)$$

Where  $W$  is optimal projection matrix, obtained by solving the generalised eigenvalue problem:

$$S_B W = \lambda S_W W \quad (10)$$

### C. Support Vector Machine

Recently, the SVM has gained much popularity in pattern classification applications due to its empirical performance. SVM belongs to the kernel methods. The kernel algorithms map data from an original space into a higher dimensional feature space using non-linear mapping. In SVM, a hyper plane used for the separation of different classes. The way in which the data separation is done by SVM is demonstrated in Fig. 4.

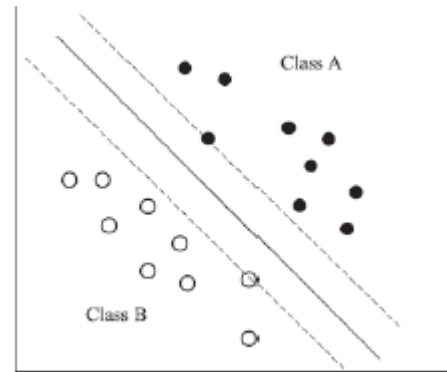


Figure 4: Data separation using SVM

SVM separates  $k$ -dimensional data using  $k-1$  dimensional hyper plane in such a way that the margin of the data sets should be maximum. In Fig. 4, the dashed lines parallel to the hyper plane contain support vectors.

## V. RESULTS

In this work, the experiments are conducted on Olivetti Research Laboratory (ORL) face database [20]. The database contain 10 different images of 40 distinct persons. The images were taken at different times, varying lighting slightly, facial expressions (open/closed eyes, smiling/non-smiling) and facial details (glasses/no-glasses). All the images are taken against a dark homogeneous background and the persons are in up-right, frontal position. The files are in .PGM format. The size of each image is  $92 \times 112$ , 8-bit grey levels. Among the 10 images for each person, 6 images were taken for training and the remaining 4 images for testing.

PCA feature extraction is as follows: First, training the samples to obtain the subspace spanned by the training images. Then, map the training images into the subspace for getting feature vectors. The extracted feature vectors are the

projection vectors of the training images. In similar way, map the test images into the subspace for getting the feature vectors of the test images. In this experiments, largest 100 eigenvectors are considered. This is to eliminate some of the eigenvectors having small eigen values, that contribute less variance in the data. The number of eigenvectors should be sufficient to separate the classes used in training. The extracted PCA features are then given for LDA feature extraction. LDA selects a set of basis vectors having smallest within-class scatter SW value and highest between- class scatter values. The subspace W is obtained by minimising the Fisher criterion.

Experiments are conducted by varying number of PCA eigen vectors and LDA eigen vectors. Repeated the experiment by extracting 100, 80, 60, 40 and 20 eigen faces. considered 10 subjects for experiments. For each value of eigen faces , number of LDA significant features are also changed. Figure 5 shows some of the eigenfaces having largest eigen value.

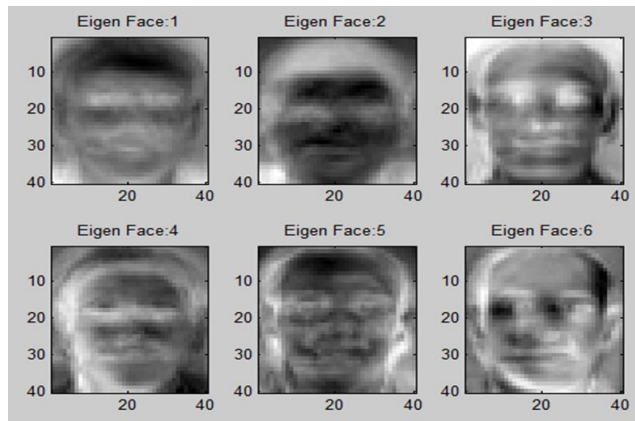


Figure 5: Eigen faces having largest eigen values

Classification accuracy of PCA+SVM and PCA+LDA+SVM is shown in Fig. 6.

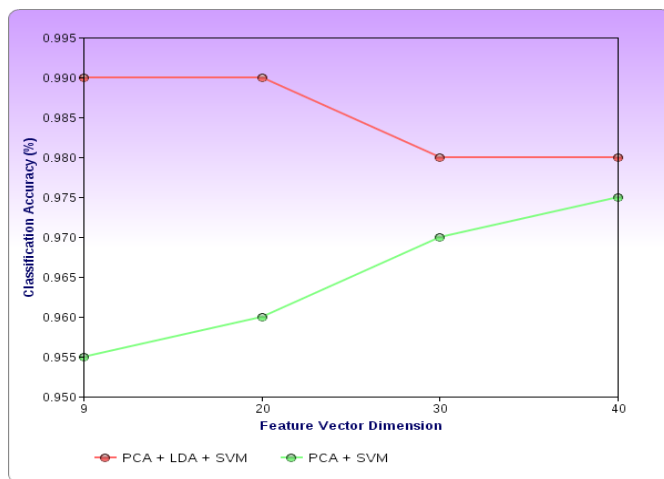


Figure 6: Feature vector dimension Vs Classification accuracy

VI CONCLUSION AND FUTURE WORK

A face retrieval system that directly depends on a face recognition system is proposed in this paper. The face recognition system is based on the combination linear dimensionality reduction techniques PCA and LDA for feature extraction and SVM for classification. Experiments are implemented on the ORL face database. The experimental results showed that, this method have improved classification accuracy. Although the classification accuracy of this work is satisfactory, it can be improved as follows: PCA is based on one dimensional vectors, so the image matrix has to be transformed into a vector prior to the feature extraction process. This is true for LDA also. If two Dimensional PCA (2DPCA) and two Dimensional LDA (2DLDA) have used instead of PCA and LDA , the image covariance matrix can directly be constructed from the original image matrices. This may reduce the computational complexity and improve the overall system performance.

REFERENCES

- [1] A. M. MartoAnez and A. C. Kak, "PCA versus LDA", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 228-233, February 2001.
- [2] S. T. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding", Science, vol 290 22, No. 5500, December 2000, pp. 2323-2326.
- [3] J. B. Tenenbaum, Vin de Silva and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction", Science, VOL 290 22, December 2000
- [4] G. Chen and S. E. Qian, "Dimensionality reduction of hyper spectral imagery using improved locally linear embedding", Journal of Applied Remote Sensing 1, 2007.
- [5] M. Pal and G. M. Foody, "Feature Selection for Classification of Hyperspectral Data by SVM", IEEE Transactions on Geoscience and Remote Sensing, Vol. 48, No. 5, May 2010
- [6] J. Li, B. Zhao and H. Zhang, "Face Recognition Based on PCA and LDA Combination Feature Extraction", The 1st International Conference on Information Science and Engineering, 2009
- [7] Md. O. Faruque and Md. Al Mehedi Hasan, "Face Recognition Using PCA and SVM", The 3rd International Conference on Anti counterfeiting, Security, and Identification in Communication, 2009.
- [8] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, NO. 7, July 1997
- [9] I. K. Fodor, "A survey of dimension reduction techniques", June 2002
- [10] S. A. Alvarez, "An exact analytical relation among recall, precision, and classification accuracy in information retrieval",
- [11] M. Turk and A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, pp. 71-86, (1991).
- [12] Olivetti Oracle Research Laboratory, The Olivetti Oracle Research Laboratory Face Database of Faces, <http://www.camorl.co.uk/facedatabase.html>

# Survey of Cloud Computing

A. Darshana Mistry

**Abstract**—Cloud computing is redefining IT operations by eliminating routine infrastructure deployment, configuration and maintenance. The cloud is a set of hardware, networks, storage, services, and interfaces that enable the delivery of computing as a service. There are three participants—end users, business management and service provider in cloud. Deployments of cloud are classified as private cloud, public cloud, community cloud and hybrid cloud. Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) are modeling service of cloud. Cloud can apply on grid computing, virtualization and paravirtualization infrastructure.

**Index Terms**—Cloud Computing, Cloud, Software as a Service, Platform as a service, Infrastructure as a service, virtualization.

## I. INTRODUCTION

CLOUD computing term refers to the delivery of scalable IT resources over the Internet (see Fig. 1) [1], as opposed to hosting and operating those resource locally, such as on a college or university network. Those resources can include applications and services, as well as infrastructure and on which they are operate.

Cloud computing represents a different way to architect and remotely manage computing resources. One has to only establish an account with Microsoft or Amazon or Google to begin building and deploying application systems into a Cloud. They can be web application that requires only http services. They might require a relation database. They might require web service infrastructure and message queues. They might be interpreted with CRM or e-commerce application services, necessitating construction of a custom technology stack to deploy into the cloud if these services are not provided there. They might require the use of new types of persistent storage they might never have to be replicated because the new storage technologies build in require reliability. They might require remote hosting and use of custom or 3<sup>rd</sup> party software systems. And they might require the capability to programmatically increase or decrease computing resource as a function of business intelligence about resource demand using virtualization.

Virtualization of computer or operating systems hides the physical characteristics of a computing platform from users; instead it shows another abstract computing platform. A hypervisor is a piece of virtualization software that allows multiple operating systems to run on a host computer concurrently. Virtualization providers include VMware, Microsoft, and Citrix systems. Virtualization is an enabler of cloud computing.

A. APCE, Gandhinagar Institute Of Technology, Gandhinagar; Computer Engineering Department, (e-mail: darshana.mistry@git.org.in).

The great advantage of cloud computing is “elasticity”: the ability to add capacity or applications almost at a moment’s notice. Companies buy exactly the amount of storage, computing power, security and other IT functions that they need from specialists in data-center computing. They get sophisticated data center services on demand, in only the amount they need and they can pay for, at service levels set with the vendor, with capabilities that can be added or subtracted at will. The metered cost, pay-as-you-go approach appeals to small and medium size enterprises; little or no capital investment and maintenance cost is needed.

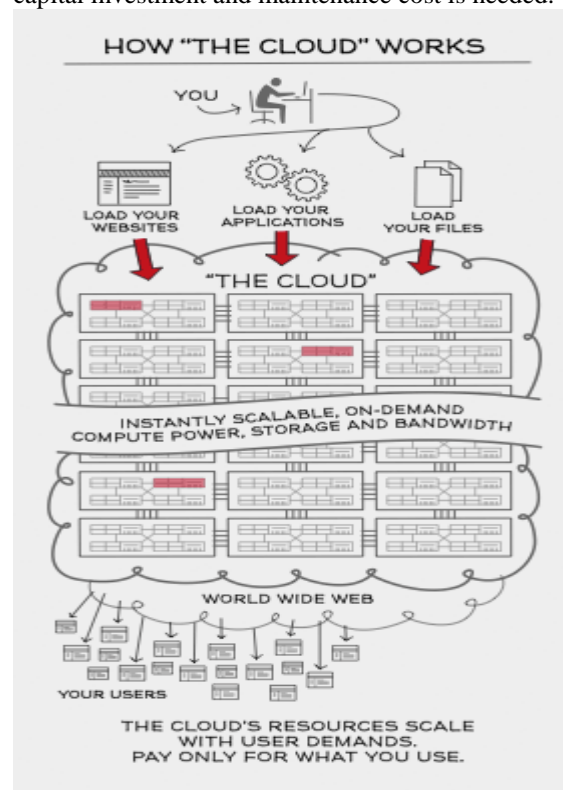


Fig.1. Cloud Computing

Drawback of cloud computing is that in the cloud, you may not have the kind of control over your data or the performance of your applications that you need, or the ability to audit or change the processes or policies under which users must work.

## II. CLOUD

The cloud itself is a set of hardware, networks, storage, services, and interfaces that enable the delivery of computing as a service. Cloud services include the delivery of software, infrastructure, and storage over the Internet based on user demand.

Participants of cloud:

- The end user doesn't really have to know anything about the underlying technology. For example, the cloud provider becomes the de facto data center.
- Business management needs to take responsibility for overall governance of data or services living in a cloud. Cloud service providers must provide a predictable and guaranteed service level and security to all their constituents.
- The Cloud service provider is responsibility for IT assets and maintenance.

Basic characteristics of cloud embodies are:

- Elasticity and the ability to scale up and down.
- Self-service provisioning and automatic deprovisioning.
- Applications programming interfaces(APIs)
- Billing and metering of service usage in a pay-as-you-go model.

#### A. Deployment Types of Clouds

Private cloud [3] are typically owned by the respective enterprise and/or leased. Functionality is not directly exposed to customer, though in some cases service with cloud enhanced features may be offered. This is similar to SaaS from the customer point of view. Example: eBay.

Public cloud (External cloud), in this resources are provisioned on fine-grained, self-service over internet, using some web service or web application, from no online-site third party providers. Example: Amazon, Google apps, Windows Azure.

Community cloud is used where many companies or organizations have alike requirements and seek for sharing infrastructure to realize cloud computing benefits. It is not as cheap as public cloud, it actually more expensive but it offers high level privacy, security and police compliance.

The environment of hybrid cloud consist of multiple external and/or internal (local) providers. By integrating services of cloud computing can have ease of transiting to public cloud and can avoid issues like PCI compliance.

#### B. Cloud Data Center

Data centers with 10,000 or more servers on site, all devoted to running very few applications that are build with consistent infrastructure components( such as racks, hardware, OS, networking and so on) as fig. 2.



Fig. 2. Cloud Data Center

Cloud data center is differing than traditional data center as below:

- In cloud data center, few applications (may be even just one) run.
- It use homogenous hardware environment and standardized management tools.
- It provides minimal application patching and updating.
- It uses single standard software architecture.
- Workload is simple.

Cost factor to build a cloud data center:

- Labor costs were 6 percent of the total costs of operating the cloud data center.
- Power distribution and cooling were 20 percent.
- Computing costs were 48 percent.

#### C. Cloud Environment Roles

The typical role distribution in Service Oriented Architecture (SOA) and in particular in Virtual Organization.

Cloud Providers offer cloud to the customer- either via dedicated APIs (PaaS), virtual machines, and/or direct access to the resources (IaaS). The hosts of cloud enhanced services (SaaS) are typically referred to as a service providers.

Cloud Resellers or Aggregators aggregate cloud platforms from cloud providers to either provide a large resource infrastructure to their customers or to provide enhanced features. This relates to community clouds in so far as the cloud aggregates may expose a single interface to a merged cloud infrastructure.

Cloud Adopters or (Software/Service) vendors enhanced their own services and capabilities by exploiting cloud platforms from cloud providers/ cloud resellers.

Cloud Customers/Users make direct use of the cloud capabilities, not to improve the services and capabilities they offer, but to make use of the direct results.

Cloud Tool Providers do not actually provide cloud capabilities, but supporting tools such as programming environments, virtual machine managements etc.

### III. MODELING SERVICES

#### A. Software as a Service (SaaS)

Business application that are hosted by the provider and delivered as a service. SaaS has its roots in an early kind of hosting operation carried out by Application Service Providers (ASPs). Customer Relationship Management (CRM) is one of the most common categories of SaaS.

Software as a Service comes in two distinct modes:

1. Simple multi-tenancy: Each customer has its own resources that are segregated from those of other customers. It amounts to a relatively inefficient form of multi-tenancy.
2. Fine grain multi-tenancy: This offers the same level of segregation but from a software engineering perspective, it's far more efficient. All resources are shared, but customer data and access capabilities are segregated within the application. This offers much superior economies of scale.

SaaS software is owned by the vendor and runs on computers in the vendor's data center. All customers of a SaaS vendors use the same software: these are one-size-fits-all solutions. Well known examples are Salesforce.com, Google's Gmail and Apps, instance messaging from AOL, Yahoo and Google, Voice over Internet Protocol (VOIP) from Vonage and Skype etc.

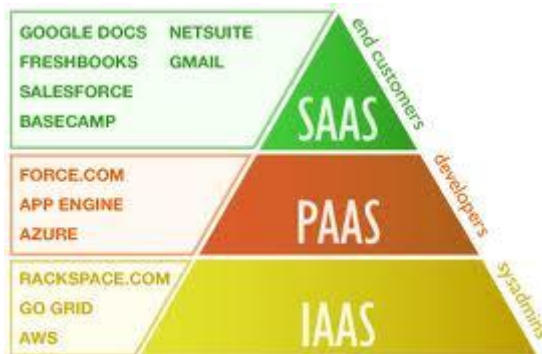


Fig. 3. Model Services

#### B. Platform as a Service (PaaS)

It delivers a solution stack—an integrated set of software that provides everything a developer needs to build an application— for both software development and runtime. PaaS can be viewed as an evolution of Web hosting. The primary benefit of PaaS is having software development and deployment capability based entirely in the cloud— hence, no management or maintenance efforts are required for the infrastructure. PaaS is inherently multi-tenant and naturally supports the whole set of Web services standards and is usually delivered with dynamic scaling.

PaaS provides virtualized servers on which user can run applications, or developed new ones, without having to worry about maintaining the operating System, server hardware, load balancing or computing capacity.

Some examples of PaaS include the Google App Engine, AppJet, Eteios, Qrimp and Force.com which is the official development environment for Salesforce.com.

#### C. The Infrastructure as a Service (IaaS)

It offers storage and compute resources that developers and IT organizations use to deliver custom business solutions. IaaS is the delivery of computer hardware (servers, networking technology, storage, and data center space) as a service. It also includes the delivery of operating systems and virtualization technology to manage the resources. In IaaS, customers rent computing resources instead of buying and installing in their own data center. The service is typically paid for on a usage basis.

Currently, the most high profile IaaS operation is Amazon's Elastic Compute Cloud (Amazon EC2). It provides a Web interface that allows customers to access virtual machines. EC2 offers scalability under the user's control with the user paying for resources by the hour.

### IV. CLOUD INFRASTRUCTURE

The infrastructure will depend on the application and how the provider has chosen to build the cloud solution [7].

#### A. Grid Computing

It applies the resources of numerous computers in a network to work a single problem at the same time. Grid computing necessitates the use of software that can divide and then sent out pieces of the program to thousands of computers. It can be done throughout the computers of an organization, or it can be done as a form of public collaboration.

Grid computing is appealing for several reasons:

- It's a cost-effective way to use a given amount of computer resources.
- It is a way to solve problems that need a tremendous amount of computing power.
- The resources of several computers can be shared cooperatively, without one computer managing the other.

#### B. Full Virtualization

Full virtualization is a technique in which a complete installation of one machine is run on another. The result is a system in which all software running on the server is within a virtual machine. For fully virtualization, specific hardware as AMD-Virtualization (AMD-V), Intel Virtualization Technology (IVT) extensions is used.

Full Virtualization has been successful for several purposes:



- Sharing a computer system among multiple users.
- Isolation users from each other and from the control program.
- Emulation hardware on another machine.

### C. Paravirtualization

It allows multiple operating system to run on a single hardware device at the same time by more efficiently using system resources, like processors and memory. In fully virtualization, entire system is emulated, but in paravirtualization, its management module operates with an operating system that had been adjusted to work in a virtual machine. Paravirtualization typically runs better than the full virtualization model, simply because in a fully virtualized deployment, all elements must be emulated.

Paravirtualization works best in these sorts of deployments:

- Disaster Recovery: in the event of catastrophe, guest instances can be moved to other hardware until the equipment can be repaired.
- Migration: Moving to new system is easier and faster because guest instances can be removed from underlying hardware.

Capacity Management: Because of easier migration, capacity management is simple to implement. It is easier to add more processing power or hard drive capacity in a virtualized environment.

## V. CLOUD FOUNDATION

Cloud Computing define as a beyond the relationship to the internet and internet technologies which hosting of hardware in an external data center (it's called as infrastructure as a service), utility component (which packages computing resources so they can be used as a utility in a always on, metered, and elastically (scalable way), platform services (middleware as a service) and application hosting (software as a service). Cloud foundations provide the basic function an application needs. These can include an underlying operating system and local support [6].

### A. Operating System

From a platform point view of view, an operating system provides a set of basic interfaces for application to use. By far the most the well known example of OS in the cloud today is Amazon's Elastic Compute Cloud (EC2). EC2 provides customer specific Linux instances running in virtual machines. From a technical perspective, it might be more accurate of EC2 as a platform for VMs rather than operating system.

Each development team is free to use whatever local support it likes in this virtual machine- Amazon does not care. The creator of one application might choose a Java EE application server and MySQL, for example, while another group might go with Ruby on Rails. EC2 customer are even free to create many Linux instances, then distributed large workloads across them in parallel, such as scientific applications. While the

services EC2 provides is quite basic, it's also very general, and so it can be used in many different ways.

### B. Local Support

In a premise platform, a developer can mix and match parts of the foundation as she sees fits. Choosing to use .Net framework on windows doesn't mandate using a particular database. Similarly, an on premise application using .NET framework is free to access the underlying Windows OS, as is an application built on a Java EE server.

The local support function in today's leading cloud foundations don't work this way. Instead, a cloud local support technology typically includes its own storage, and it hides whatever the underlying OS might be.

Cloud computing should be so attractive because of its scalability, but to make an application built on a cloud foundation handle Internet size load requires limiting in its same ways. By making local supporting functions more specialized, a cloud platform provider has more freedom to optimize the application environment. Accordingly, each set of local support functions in cloud foundations today focuses on supporting a particular kind of application. For example, Google's AppEngine provides local support for running Python Web applications. Along with a standard Python runtime, AppEngine also includes a hierarchical data store with its own query language. Microsoft also provides local support for application in the cloud as CRM live offering. Based on the Dynamics CRM platform, this technology target data oriented business applications, much like Force. Com. And like Force.com and AppEngine, it includes both run time application support and a data store.

## VI. CONCLUSION

Cloud computing is work on virtualization, so it hides physical characteristics of a computer platform from users. Cloud computing is more popular because of ability to add capacity or application (elasticity) at any moment, pay as you go approach, no capital investment and maintenance cost. Cloud should be uniquely identifiable, dynamic configuration, and secure. Clouds are used for three main purpose Software as a Service, Platform as a Service, Infrastructure as a Service.

## REFERENCES

- [1] "7 things should you know about cloud computing", [www.educause.edu/.../7ThingsYouShouldKnowAboutCloud/176856](http://www.educause.edu/.../7ThingsYouShouldKnowAboutCloud/176856)
- [2] T. B. Winans and J. S. Brown, "Cloud Computing", A collectin of working papers, May 2009.
- [3] J. Hurwitz, R. Bloor, M. Kaurfman, F. Halper, "Cloud Computing for Dummies" Hand Book, Will -India
- [4] K. Jeffery, B. Neidecker, "The Future of Cloud Computing", Opportunity for Europeon Cloud Computing Beyond 2010, verion 1.0.
- [5] IA newsletter: the Newsletter for Information Assurance Technology Professions, Volume 13, number 2, Spring 2010.
- [6] G. Boss, P. Malladi, D. Quan, L. Legrengi, "Cloud Computing", High Performance on Demand Solutions, Version 1, October 2007.
- [7] D. Chappell, "A short introduction to cloud platforms-an enterprise oriented view", August 2008
- [8] "Cloud Computing Basics", [www.mhprofessional.com/.../0071626948\\_chap01.pdf](http://www.mhprofessional.com/.../0071626948_chap01.pdf) - United States,

# Survey of Chaotic Maps for Image Security/ Encryption

A. Lokesh Pitambar Gagnani

**Abstract**—Cryptography is about communication in the presence of an adversary. It encompasses many problems like encryption, authentication, and key distribution to name a few. The field of modern cryptography provides a theoretical foundation based on which one can understand what exactly these problems are, how to evaluate protocols that purport to solve them and how to build protocols in whose security one can have confidence.

Advanced digital technologies have made multimedia data widely available. Recently, multimedia applications become common in practice and thus security of multimedia data has become main concern.

The basic issues pertaining to the problem of encryption has been discussed and also a survey on image encryption techniques based on chaotic schemes has been dealt in the present communication. Chaotic Encryption Method seems to be much better than traditional encryption methods used today. Chaotic encryption is the new direction of cryptography. It makes use of chaotic system properties such as sensitive to initial condition and loss of information. Many chaos-based encryption methods have been presented and discussed in the last two decades. In order to reach higher performance, these methods take advantage of the more and more complex behavior of chaotic signals. This paper contributes by comparing the past chaotic maps used for image encryption.

**Index Terms**—Chaotic Encryption, Chaotic Map, Cryptography, Image Encryption, Security.

## I. INTRODUCTION

Cryptography is the science of protecting the privacy of information during communication under hostile conditions. In the present era of information technology and proliferating computer network communications, cryptography assumes special importance. Cryptography is now routinely used to protect data, which must be communicated and/or saved over long periods, to protect electronic fund transfers and classified communications.

Current cryptographic techniques are based on number theoretic or algebraic concepts. Chaos is another paradigm,

A. Pursuing Masters in IT field from Shantilal Shah Engineering College, Bhavnagar. ( e-mail: lokesh\_jolly05@yahoo.co.in)

which seems promising. Chaos is an offshoot from the field of nonlinear dynamics and has been widely studied. A large

number of applications in real systems, both man-made and natural, are being investigated using this novel approach of nonlinear dynamics. The chaotic behavior is a subtle behavior of a nonlinear system, which apparently looks random. However, this randomness has no stochastic origin. It is purely resulting from the defining deterministic processes. The important characteristics of chaos are its extreme sensitivity to initial conditions of the system.

## II. SECURITY ISSUE

Multimedia security in general is provided by a method or a set of methods used to protect the multimedia content. These methods are heavily based on cryptography and they enable either communication security, or security against piracy (Digital Rights Management and watermarking), or both. Communication security of digital images and textual digital media can be accomplished by means of standard symmetric key cryptography. Such media can be treated as binary sequence and the whole data can be encrypted using a cryptosystem such as Advanced Encryption Standard (AES) or Data Encryption Standard (DES). In general, when the multimedia data is static (not a real-time streaming) it can be treated as a regular binary data and the conventional encryption techniques can be used. Deciding upon what level of security is needed is harder than it looks. To identify an optimal security level, the cost of the multimedia information to be protected and the cost of the protection itself are to be compared carefully. At present, there are many available image encryption algorithms such as Arnold map, Tangram algorithm, Baker's transformation, Magic cube transformation, and Affine transformation, etc . Chaotic Map include Arnold's Cat map, Baker's map, Henon map, Ikeda map, Logistic map, Lorenz attractor, Tent map, Tinkerbell map. In some algorithms, the secret-key and algorithm cannot be separated effectively. This does not satisfy the requirements of the modern cryptographic mechanism and are prone to various attacks. In recent years, the image encryption has been developed to overcome above disadvantages.

III. CHAOTIC SYSTEM

A chaotic system is one in which a tiny change can have a huge effect. Chaos theory is a scientific discipline that focuses on the study of nonlinear systems that are highly sensitive to initial conditions that is similar to random behavior, and continuous system. The properties of chaotic systems are:

- (i) Deterministic, this means that they have some determining mathematical equations ruling their behavior.
- (ii) Unpredictable and non-linear, this means they are sensitive to initial conditions. Even a very slight change in the starting point can lead to significant different outcomes.
- (iii) Appear to be random and disorderly but in actual fact they are not. Beneath the random behavior there is a sense of order and pattern. The highly unpredictable and random-look nature of chaotic output is the most attractive feature of deterministic chaotic system that may lead to various novel applications.

IV. CRYPTOGRAPHY SYSTEMS AND CHAOTIC SYSTEMS

In cryptographic terminology, *confusion* refers to the process of substitution. The confusion/substitution is intended to make the relationship between the key and the ciphertext as complex as possible. In cryptography, diffusion refers to the process of rearranging or spreading out the bits in the message so that redundancy in the plaintext is spread out over the complete ciphertext. Particularly in images, where the redundancy is large, the process of diffusion is a necessary requirement to develop a secure encryption technique. In image encryption, the diffusion process changes the statistical properties of the plain-image by spreading the influence of each bit of the plain-image over the whole cipher-image, which removes the possibility of differential attacks by comparing the pair of plain and cipher-images.

The difference between these systems is depicted in Table I:

TABLE I  
CRYPTOGRAPHY Vs CHAOS

Chaos	Cryptography
Sensitivity to initial conditions.	Diffusion of a small change in plaintext
Sensitivity to control parameters	Confusion of a small change in secret keys.
Mixing property: spreading out the influence of a single plaintext digit over many cipher text digit.	Avalanche Effect: small change in either the plaintext or the key should produce a significant change in the cipher text.
Deterministic dynamics	Deterministic pseudo randomness.

V. CHAOTIC MAPS

A chaotic map is a map (=evolution function) that exhibits some sort of chaotic behavior. Maps may be parameterized by a discrete-time or a continuous-time parameter. Discrete maps usually take the form of iterated functions. Chaotic maps often

occur in the study of dynamical systems. List of Chaotic Maps commonly used in Image Encryption are as follows:

A. Arnold Cat Map (ACM):

Arnold mapping is a typical chaotic mapping.

Let  $X = \begin{bmatrix} x \\ y \end{bmatrix}$  be the  $n \times n$  matrix, then the ACM transformation

$$\Gamma: \begin{bmatrix} x' \\ y' \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \pmod n$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \pmod n$$

$$\Gamma: (x', y') \rightarrow (x + y, x + 2y) \pmod n$$

Where  $x'$  and  $y'$  are new co-ordinates &  $x$  and  $y$  are original co-ordinates of image. Also  $n$  is the size of image & we consider only square images.

B. Swarm Intelligence

Swarm Intelligence is a design framework based on social insect behavior. Social insects such as ants, bees, and wasps are unique in the way these simple individuals cooperate to accomplish complex, difficult tasks. This cooperation is distributed among the entire population, without any centralized control or the provision of a global model. In particle swarm model each particle adjusts its position according to its own experience, and according to the experience of neighbors, making use of the best position encountered by itself and its neighbors. The particle swarm model consists of a swarm of particles, moving iteratively through the  $m$ -dimension problem space to search the new solutions (position).

C. Lorenz Attractor/System

The Lorenz attractor, named for Edward N. Lorenz, is an example of a non-linear dynamic system corresponding to the long-term behaviour of the Lorenz oscillator. The Lorenz oscillator is a 3-dimensional dynamical system that exhibits chaotic flow, noted for its figure eight shape. The map shows how the state of a dynamical system (the three variables of a three-dimensional system) evolves over time in a complex, non-repeating pattern.

The equations that govern the Lorenz oscillator are:

$$x' = \sigma(y - x)$$

$$y' = x(\rho - z) - y$$

$$z' = xy - \beta z$$

where  $\sigma$  is called the Prandtl number and  $\rho$  is called the Rayleigh number. All  $\sigma, \rho, \beta > 0$ , but usually  $\sigma = 10, \beta = 8 / 3$  and  $\rho$  is varied. The system exhibits chaotic behavior for  $\rho = 28$  but displays knotted periodic orbits for other values of  $\rho$ .

The Lorenz attractor generated with above values is shown in Fig 1:

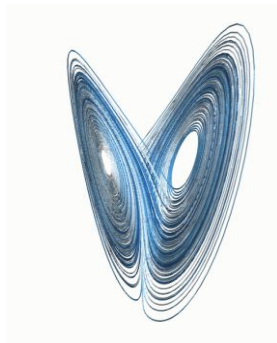


Figure 1: Lorenz Attractor

#### D. Chen System

Chen System is similar to the Lorenz attractor chaos system. It is described by following set of equations:

$$x' = a(y - x)$$

$$y' = x(c - a) - xz + cy$$

$$z' = xy - bz$$

when  $a=35$ ,  $b=3$  and  $c=28$  the system exhibits chaotic behavior.

The Chen attractor generated with above values is shown in Figure 2:

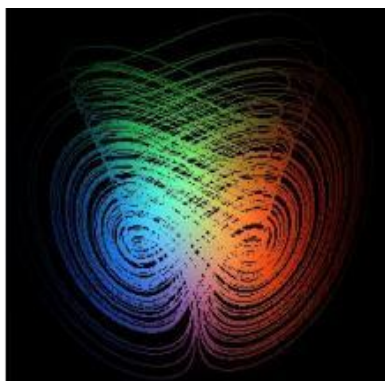


Figure 2: Chen Attractor

#### E. Lu Map

The Lu chaotic map is a 3D chaotic map; it is described by following set of equations:

$$x' = a(y - x)$$

$$y' = -xz + cy$$

$$z' = xy - bz$$

where  $(x', y', z')$  is a system trace,  $(a, b, c)$  is system parameters. When  $a=36$ ,  $b=3$ ,  $c=20$ , the system contain a strange attractor and being in chaotic state. Giving the initial value  $x_0, y_0, z_0$  of the Lu map, let the system iterate  $N \times N$  times, produce three sequence values each time. Then these

sequences have the same characteristics of chaos signals namely the characteristics of randomness, ergodic, and the sensibility to initial value, so they can be used on image encryption.

#### F. Logistic Chaotic Map

Logistic mapping is a simple and widely studied one dimensional discrete nonlinear dynamic system. It's defined as follows:

$$x_{n+1} = \mu x_n (1 - x_n)$$

Among that  $0 \leq \mu \leq 4$  is known as the branch parameter,  $x_n$  belongs to  $(0,1)$ ,  $n = 0, 1, 2, 3, \dots$ , known as the state. At the time of  $3.5699456 \dots \leq \mu \leq 4$ , Logistic mapping is in a chaotic state, this time it generates a sequence  $\{x_n\}$ ,  $n = 0, 1, 2, 3, \dots$ , which is non-periodic, non-convergence, and is very sensitive to initial values.

#### G. Henon Map

Henon is a simple non-trivial two-dimensional mapping. Its mathematical expression as follows:

$$X_{n+1} = 1 - aX_n^2 + bY_n$$

$$Y_{n+1} = X_n$$

In order to fit the characteristics of image encryption, image encryption algorithm always do the Henon mapping modulo transformation. After that, the Henon mapping as follows:

$$X_{n+1} = 1 - aX_n^2 + bY_n \text{ mod } G$$

$$Y_{n+1} = X_n \text{ mod } G$$

$G$  is the maximum of pixel value, for example, if it is the gray value and to use a byte, then  $G=256$ .

#### VI CONCLUSION

A conclusion is obtained that chaos systems are more powerful than the cryptographic systems due to the property of sensitivity to initial parameters. Also a survey of various commonly used chaotic maps was done. The chaotic maps prove useful for image security specially in image encryption filed.

#### REFERENCES

- [1] Reda M Hussein, Hatem S Ahmed, Wail F Abd El-Wahed, "New Encryption Scheme on Swarm Intelligence Chaotic Map", IEEE Society.
- [2] Ling Bin, Liu Lichen & Zhang Jan, "Image Encryption algorithm based on chaotic map & S-DES", pp. 41-44, 2010, IEEE.
- [3] Yuanmei Wang, Tao Li, "Study of Image Encryption Algorithm based on Arnold Transformation & Chaotic System", pp. 449-451, 2010, IEEE Society.
- [4] Senthil Arumugam, Kiruba Jothi, "Image Encryption Algorithm based on Improved 3d Chaotic Map", 2010, IEEE Society.
- [5] Weizhong Yu, Guoqiang Bai, "A New High Performance Image Encryption based on Chua's Attractor", pp.874-878, 2010, IEEE Society.

- [6] A New High Performance Image Encryption based on Chua's Attractor", pp.874-878, 2010, IEEE Society.
- [7] Tianshou Zhou, Yun Tang and Guanrong Chen, "Chen's Attractor Exists ", pp. 3167-3177, 2004, International Journal of Bifurcation & Chaos, Vol. 14, No. 9.
- [8] Jinhua Lu, Tianshou Zhou, Guanrong Chen and Suochin Zhang, "The Compound Structure of Chen's Attractor ", pp. 855-858, 2002, International Journal of Bifurcation & Chaos, Vol. 12, No. 4.
- [9] Jinhua Lu, Guanrong Chen, "A new Chaotic attractor coined ", pp. 659-661, 2002, International Journal of Bifurcation & Chaos, Vol. 12, No. 3.
- [10] Matthew Carriolo, Brown University, "The Lorenz attractor, chaos and fluid flow ", 2005.

# “BAD LOGIN” as a Mechanism For Secure Code Execution and Behavior Analysis

A. Patel Sudhir, B. Dr. Nilesh K Modi

**Abstract:** - In this paper we present an overview of the fundamentals of user authentication based on hand internet security and trial error method. This paper presents some new features for Internet security and authentication. We will introduce the new algorithm for login system and authentication for commodity user.

## I. INTRODUCTION

Today, businesses and individuals are entrusting progressively greater amounts of security-sensitive data to computers, both their own and those of third parties. To be worthy of this trust, these computers must ensure that the data is handled with care (e.g., as the user expects), and protected from external threats. Unfortunately, today's computer platforms provide little assurance on either front. Most platforms still run code designed primarily for features, not security. While it is difficult to measure precisely, multiple heuristics suggest that code quality has improved relatively little with respect to security. For example, the majority of coding projects on Source Forge employ type-unsafe languages (e.g., C or C++)<sup>[1]</sup>. The National Vulnerabilities Database<sup>[2]</sup> catalogues thousands of new software vulnerability reports each year, and recent studies indicate that over 25% of US computers are infected with some form of malicious software<sup>[3]</sup>.

These vulnerabilities are particularly troubling when coupled with the increasingly hostile environment to which computers (and users) are exposed. Indeed, the lucrative and difficult-to-prosecute crimes that computers facilitate have given rise to a burgeoning criminal underground in which sophisticated, Financially-motivated attackers collaborate to monetize exploited computers and stolen user data<sup>[4]</sup>. These ne'er-do-wells can employ automated, turn-key.

A. MCA, Ph.D. (Pursuing) Mewar University, Chhittogarh. (Raj)Asst. Prof & Head - D.L.Patel Inst. of Mgt & Tech. MCA College, Himatnagar, Gujarat, India [patel\\_sudhirkumar@gtu.edu.in](mailto:patel_sudhirkumar@gtu.edu.in), [sudhirkumar.patel@gmail.com](mailto:sudhirkumar.patel@gmail.com)

B. Professor & Head, S.V. I.C.S, MCA College, KADI, Gujarat [drnileshmodi@yahoo.com](mailto:drnileshmodi@yahoo.com)

packages to attack thousands of potential victims every second<sup>[5]</sup>. The victim computers are then often formed into coordinated “botnets” of tens or hundreds of thousands of machines and used to send spam or launch Distributed Denial of Service (DDoS) attacks<sup>[6,7]</sup>.

As a result, from the moment a user digitizes her data, it is under constant assault from all sides. Malicious software on the user's computer may snatch up the user's private data and ship it to foreign lands. If the user entrusts her data or computations to a remote service, then the remote computers may be subject to all manner of physical attacks, as surveyed in Section. Furthermore, experience demonstrates that remote workers will attempt to return forged results even when the only payoff is an improvement of their standings in an online ranking<sup>[8]</sup>; when there is a potential to profit from such skullduggery, workers' temptations can only increase.

Sadly, these attacks on user data succeed all too often. They contribute to the over 3.6 million U.S. households that were victims of identity in the year 2004 alone<sup>[9]</sup>. They also undermine users' trust in electronic systems, and hence inhibit both current and future systems. For example, a 2005 Consumer Reports survey found that 29% of consumers had cut back on – and 25% had stopped – shopping online due to fears of fraud and identity theft<sup>[10]</sup>.

The following figure shows the authentication process in existing system.

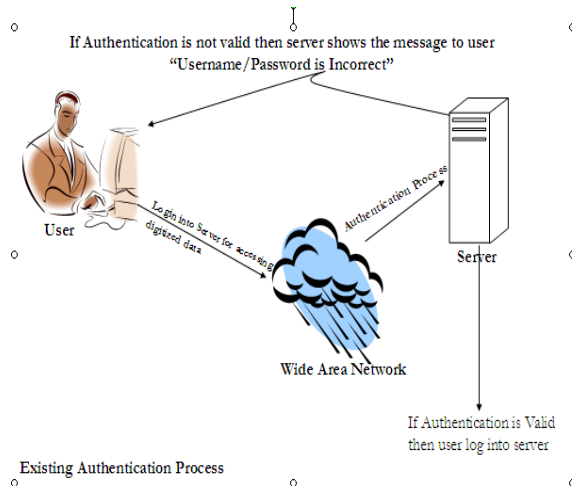
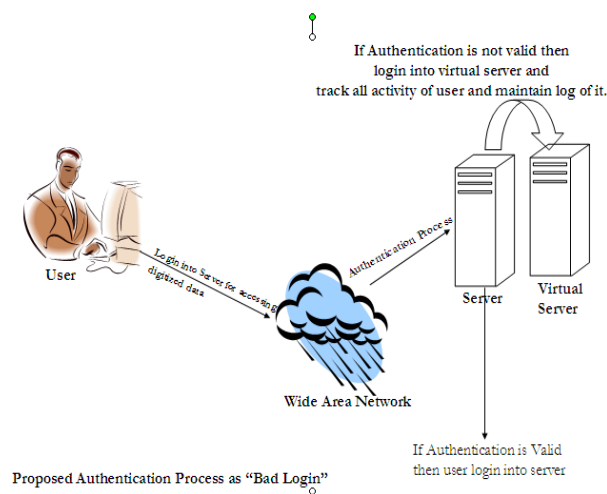


Fig-1 shows the normal login system for any user. In normal login system user use username and password. If password or user id is wrong then system show the message to user “you username or password is wrong” and redirect to login page so user enter valid username and password. In this scenario may hacker immediately come to know password is not valid or valid and he try different password with same user-id and hack the user account based on trial and error method.



**Fig 2 “Proposed Authentication System”**

As show in fig-2 we are implement new authentication system where normal user will track the hacker activity. In our proposed authentication system, user enter user-id and password if password is not valid then we log

user into virtual system and track the all activity of that user and maintain the log of it. During the next valid login user will prompt with last bad login activity log so user come to know someone try to do un-authentication work in his system.

Bad login is concept which is used for developing trust platform module (T.P.M) on user local computer and secures the digital data of user.

### Leveraging Secure Code Execution to Improve Network Protocols

If we can provide secure code execution on endhosts, the next frontier is to examine how such trust can be used to improve the performance and efficiency of network applications. In other words, if endhosts (or at least portions of each endhost) can be trusted, then network infrastructure no longer needs to arduously and imprecisely reconstruct data already known by the endhosts.

Through the design of a general-purpose architecture we call Assayer<sup>[11]</sup>, we explore the issues in providing trusted host-based data, including the balance between useful information and user privacy, and the tradeoffs between security and efficiency. We also evaluate the usefulness of such information in three case studies: spam identification, distributed denial-of-service attack mitigation, and super-spreader worm detection.

To gain insight into the performance we could expect from such a system, we implement and evaluate a basic Assayer prototype. Our prototype requires fewer than 1000 lines of code on the endhost. Endhosts can annotate their outbound traffic in a few microseconds, and these annotations can be checked efficiently; even packet-level annotations on a gigabit link can be checked with a loss in throughput of only 3.7-18.3 %, depending on packet size.

### Secure Code Execution despite Untrusted Software and Hardware

With Flicker, we assume that the user’s computer is physically secure. To generalize Flicker’s results, we need techniques to establish trust in code execution when even the hardware is untrustworthy. This scenario is particularly compelling as the growth of “cloud computing” and the proliferation of mobile devices

contribute to the desire to outsource computing from a client device to an online service. In these applications, how can the client be assured that the secrecy of her data will be protected? Equally importantly, how can the client verify that the result returned is correct, without redoing the computation?

While various forms of homomorphic encryption can provide data secrecy<sup>[12,13]</sup>, the results in demonstrate that we can efficiently verify the results of arbitrary tasks (abstracted as function evaluations) on a computational service (e.g., in the cloud) without trusting any hardware or software on that system. This contrasts with previous approaches that were inefficient or that could only verify the results of restricted function families. To formalize secure computational outsourcing, introduces the notion of verifiable computing<sup>[14]</sup>. Abstractly, a client wishes to evaluate a function  $F$  (e.g., sign a document or manipulate a photograph) over various, dynamically selected inputs  $x_1, \dots, x_k$  on one or more untrusted computers, and then verify that the values returned are indeed the result of applying  $F$  to the given inputs. The critical requirement, which precludes the use of previous solutions, is that the client's effort to generate and verify work instances must be substantially less than that required to perform the computation on her own. Drawing on techniques from multi-party secure computation, as well as some recent developments in lattice-based cryptography, we present the first protocol for verifiable computing. It provably provides computational integrity for work done by an untrusted party; it also provides provable secrecy for the computation's inputs and outputs. Moreover, the protocol provides asymptotically optimal performance (amortized over multiple inputs). Specifically, the protocol requires a one-time pre-processing stage which takes  $O(|C|)$  time, where  $C$  is the smallest known Boolean circuit computing  $F$ . For each work instance, the client performs  $O(|m|)$  work to prepare an  $m$ -bit input, the worker performs  $O(|C|)$  work to compute the results, and the client performs  $O(|n|)$  work to verify the  $n$ -bit result.

This result shows that we can outsource arbitrary computations to untrusted workers, preserve the secrecy of the data, and efficiently verify that the computations were done correctly. Thus, verifiable computing could be used, for instance, in a distributed computing project

like Folding@home<sup>[15]</sup>, which outsources the simulation of protein folding to millions of Internet users. To prevent cheating, these projects often assign the same work unit to multiple clients and compare the results; verifiable computing would eliminate these redundant computations and provide strong cryptographic protections against colluding workers.

Thus, even without secure hardware, these results show that we can leverage a user's trust in one device to verify (and hence trust) the results of computations performed by an arbitrary number of remote, untrusted commodity computers.

## II. OBJECTIVE AND SCOPE

### A Vision for a Better World

I envision a future in which average computer users can easily and securely use their computers to perform sensitive tasks (e.g., paying bills, shopping online, or accessing medical records), while still retaining the flexibility and performance expected of modern computers. Users will regard a computer attack not as a disaster that empties bank accounts or destroys documents, but as a minor annoyance; like a blown electrical fuse, it will be easy to detect and simple to remedy. Providing security as a largely invisible default will allow the information age to finally reach its true potential: users will submit their personal information to online sites, not blindly or with trepidation, but with confidence that it cannot be stolen or misused; businesses and consumers will feel perfectly secure outsourcing work to computational services; and remote, web-based applications will provide the same level of privacy, security, and availability that native applications do.

Achieving this goal will require advances on many fronts: better programming languages, better operating systems, better network protocols, and better definitions of security. More fundamentally, however, we must enable both computers and users to make accurate, informed trust decisions. After all, even if software does improve, we must be able to determine which systems employ the new and improved software! This applies both to users and to network components. In other words, it is critical that a user be able to judge whether a system (either local or remote) should be trusted before she hands over her sensitive data. Similarly, if a network element (e.g., a router) can trust information from an end host, then numerous protocol optimizations become possible.



In this work, we focus on a definition of trust similar to the definition of a Nash Equilibrium; for example, to trust an entity X with her private data (or with a security-sensitive task), a user Alice must believe that at no point in the future will she have cause to regret having given her data (or entrusted her task) to X. As a result, this dissertation examines techniques that provide firm evidence on which to base such a belief. As an additional constraint, we concentrate on average users and commodity systems, rather than on advanced users, special-purpose computers, or highly constrained environments (such as those found within the military).

Alas, previous efforts to construct trustworthy systems “from the ground up” have proven difficult, time-consuming, and unable to keep pace with the changing demands of the marketplace. For example, the VAX VMM security kernel was developed over the course of eight years of considerable effort, but in the end, the project failed, and the kernel was never deployed. This failure was due, in part, to the absence of support for Ethernet – an emerging and highly popular feature considered critical by the time the kernel was completed, but not anticipated when it was initially designed. Thus, such efforts have typically been doomed, and their methods have not been adopted into the mainstream of software development.

The concept of bad login is about to prevent the hacker to hack digital data through the Trail & Error method.

Today we all know that, Internet user increasing day by day but they do not aware about hacker activity.

People who generally not aware about HACKER they will set weak password like birth-date, mobile-number, mom-dad name, school name etc. which is belong to user profile and easy available on social engineering website. Hacker will easily track the user activity by getting his password via some mechanism we called as HACKING.

Sometime HACKER use trial and error method, where Hacker enter different password with same user-id. PASSWORD entered by HACKER is relative information of user (birth-date, mobile-number, mom-dad name, school name etc).

### **Bootstrapping Trust in a Commodity Computer**

Initially, we focus on the problem of allowing a user to bootstrap trust in her own personal computer. This problem is fundamental, common, and should be easier

than other potential scenarios. In other words, if we cannot establish trust in the user’s computer, we are unlikely to be able to establish trust in a remote computer. When working with her own computer, the user can at least be reasonably certain that the computer is physically secure; i.e., an attacker has not tampered with the computer’s hardware configuration. Such an assumption aligns quite naturally with standard human intuition about security: a resource (e.g., a physical key) that an individual physically controls is typically more secure than a resource she gives to someone else. Fortunately, the physical protection of valuable items has been a major focus of human ingenuity over the past several millennia.

If the user’s computer is physically secure, then we can make use of special-purpose hardware to support the user’s security decisions. While a full-blown secure coprocessor, such as the IBM 4758, might be appealing, cost and performance considerations make deployment difficult. However, for the last few years, many commodity computers have come equipped with a Bad login Trusted Platform Module (TPM) that can be used for a variety of security-related purposes.

Unfortunately, at present, no standard mechanism exists for establishing trust in the TPM on a local machine. Indeed, any straightforward approach falls victim to a cuckoo attack. In this attack, the adversary extracts the private keys from a TPM under his physical control. These keys can be given to malicious software present on the user’s local computer, in order to fool the user into accepting reassurances from the adversary’s TPM, rather than her own.

We propose a formal model for establishing trust in a platform. The model reveals the cuckoo attack problem and suggests potential solutions. We survey the usability challenges entailed by each solution, and suggest preferred approaches to enable a user to bootstrap trust in the secure hardware on her personal computer.

### **III. IMPLEMENTATION:**

In this phase the proposed model will be implemented to achieve the objective of the system so that the model can be referred to stop the hacker attack (**Bad Login**). This phase use the document from the design phase and requirement document from the analysis phase to obtain and use the model.

### **IV. CONCLUSIONS:**

Motivated by the trend of entrusting sensitive data and services to insecure computers, we develop techniques

that allow a user to extend her trust in one device in order to securely utilize another device or service. A key constraint is our focus on commodity computers, particularly the need to preserve the performance and features expected of such platforms. Using a logical framework, we analyze the perils of establishing trust in a local computer equipped with commodity (i.e., low-cost) security hardware and provide guidance on selecting a mechanism that preserves security while minimizing changes to existing computer designs. To make the results of such an interaction meaningful, we develop the Flicker architecture for providing an on-demand, secure execution environment for security-sensitive code. Building on recent improvements in commodity CPUs, Flicker provides strong isolation, reporting, and state preservation for security sensitive code. Because it runs only on demand, Flicker imposes zero overhead on non-security-sensitive code.

## V. BOOK REFERENCE

1. J. Howell, J. R. Douceur, J. Elson, and J. R. Lorch. Leveraging legacy code to deploy desktop applications on the web. In Proceedings of the USENIX Symposium on Operating Systems Design and Implementation (OSDI), Dec.2008 (Referenced on page 14)
2. NIST Computer Security Resource Center (CSRC). National vulnerability database. <http://nvd.nist.gov/home.cfm>.
3. Organization for Economic Co-operation and Development. Malicious software (malware): a security threat to the internet economy. <http://www.oecd.org/dataoecd/53/34/40724457.pdf>. (Referenced on page 14)
4. J. Franklin, V. Paxson, A. Perrig, and S. Savage. An inquiry into the nature and causes of the wealth of internet miscreants. In Proceedings of the ACM Conference on Computer and Communications Security (CCS).
5. Sophos. Do-it-yourself phishing kits found on the Internet, reveals <http://www.sophos.com/spaminfo/articles/diyphishing.html>.
6. T.Holz, M. Steiner, F. Dahl, E. Biersack, and F. Freiling. Measurements and mitigation of peer-to-peer-based botnets: A case study on storm worm. In Proceedings of the USENIX Workshop on Large-Scale Exploits and Emergent threats (LEET),
7. E. Messmer. Downadup/conflicker worm: When will the next shoe fall? Network World.
8. D. Molnar. \_e SETI@Home problem. ACM Crossroads,. (Referenced on pages 15, 92 and 158.)
9. Department of Justice, Bureau of Statistics. Press release: Identity theft !. <http://bjs.ojp.usdoj.gov/content/pub/press/it04pr.cfm>, (Referenced on page 15.)
10. Princeton Survey Research Associates International. Leap of Faith: Using the Internet Despite the Dangers (Results of a National Survey of Internet Users for Consumer Reports Web Watch). <http://www.consumerwebwatch.org/pdfs/princeton.pdf>,
11. B. Parno, Z. Zhou, and A. Perrig. Help me help you: Using trustworthy host-based information in the network. Technical Report CMU-CyLab-09-016, Carnegie Mellon University, Cylab, Nov. 2009. (Referenced on page 19.)
12. C. Gentry. Fully homomorphic encryption using ideal lattices. In Proceedings of the ACM Symposium on the theory of Computing (STOC), (Referenced on pages 20, 50, 160, 161, 165, and 166.)
13. M. van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan. Fully homomorphic encryption over the integers. In Proceedings of EuroCrypt, June 2010. (Referenced on pages 20 and 50.)
14. R. Gennaro, C. Gentry, and B. Parno. Non-interactive verifiable computation: Outsourcing computation to untrusted workers. In Proceedings of CRYPTO, Aug. 2010. (Referenced on page 20)
15. Pande Lab. The folding@home project. Stanford University, <http://folding.stanford.edu/>. (Referenced on pages 21,74,158,162 and 181)

# Palm Print Based Biometric Authentication System

A.Ms. Swati Verma , B. Mr.Abhishek Misal

**Abstract-** Palmprint based personal authentication has gained preference over other biometric modalities due to its ease of acquisition , high user acceptance and reliability. In this paper we present an overview of the fundamentals of personal authentication based on hand geometry measurements and palm print features. This paper presents some new features for the palm print based authentication. The Region of interest (ROI) is extracted from the palm print image by finding a tangent to the curves between fingers. The perpendicular bisector of this tangent and the tangent itself help demarcate the rectangular area that forms the ROI of the palm print

**Key Word:** Local Binary Pattern, Support Vector Machines.

## I. INTRODUCTION

The palmprint as a biometric modality is slowly but decisively gaining acceptance in the field of biometrics. As compared to other biometric modalities, it is bestowed with enormous information that is a boon of its discriminating power. We will discuss a few important contributions made on this modality. A palmprint authentication system in [1] is specifically designed to overcome the limitations of the contemporary biometric authentication systems. In this system, geometric and pseudo Zernike moments are employed as feature extractors from the decomposed palmprint image. Before moment computation, wavelet transform is applied to decompose a palmprint image into lower and higher frequency sub bands. This decomposition reduces the computational burden of the moment calculations drastically. The generated wavelet moments based features are used to create a cancelable verification key with a set of random data. This private binary key can be canceled and replaced. Besides this key also possesses high data capture offset tolerance, with highly correlated bit strings for intra-class population. This property allows a clear separation of the genuine and imposter populations as well as attains the zero Equal Error Rate attainment which cannot be realized by the conventional biometric based authentication system. Matching of palmprints in [2] deals with the feasibility of person identification based on a set of feature points extracted along the prominent palm lines (and the associated line orientation) from a given palmprint image. Next a decision is made whether two palmprints belong to the same hand by computing a matching score between the corresponding sets of feature points of the

two palmprints. Two sets of features/orientations are matched using point matching technique which takes into account the nonlinear deformations as well as the outliers present in the two sets. The estimates of the matching score distributions for the genuine and imposter sets of palm pairs indicate that palmprints have a good discrimination power. Palmprint Verification with moments in [3] introduces an experimental evaluation of the effectiveness of utilizing three well known orthogonal moments, namely Zernike moments, pseudo Zernike moments and Legendre moments for palmprint verification. The idea of implementing orthogonal moments as palmprint feature extractors is prompted by the fact that the principal features of palmprint are based on the line structure. These orthogonal moments are able to define statistical and geometrical features containing line structure information about palmprint. Experimental results show that the performance of the system is dependent on the moment order as well as the type of moments. The orthogonal property of these moments is able to characterize independent features of the palmprint image and thus have minimum information redundancy in a moment set. Palmprint verification using complex wavelet transform [4] moots a modified complex wavelet structural similarity index (CW-SSIM) as the matching score for identifying the input palmprint. The local structure information of a palmprint is hidden in the relative phase patterns of the complex wavelet coefficients and a constant phase shift of all coefficients does not change the structure of local image features. Since CW-SSIM is robust to translation, small rotation and distortion, a fast rough alignment of palmprint images is sufficient. CWSSIM is also in sensitive to luminance and contrast changes. Palmprint classification in [5] is an important indexing mechanism in a palmprint database. This algorithm uses a novel representation and is based on two-stage classifier that provides even-distributed categories[6]. A novel representation scheme is directly derived from the principal line structures The representation does not use wrinkles, and singular points. It is capable of tolerating poor image quality. In this paper, we will explore feature extraction using Local Binary Pattern Operator, fuzzy features and wavelet features extracted from a palmprint. Though Gabor features have been extensively used on palmprints, it is observed that they are sensitive to slight changes in size and orientation of a palm. An effort is made to

work on two counts: An efficient extraction of Region of Interest (ROI) and an effective features election[7,16].

## II. ROI EXTRACTION

### A: The Preprocessing for ROI extraction

The following steps outline the ROI extraction from a palmprint:

1. Take the original image (See Fig. 1) and convert it from RGB to gray scale.
2. Crop a fixed section of the image not touching the glass (See Fig. 2).
3. Rotate this section based on the type of hand– left or right so that the image is in the specified direction.
4. Find the histogram of the image.
5. Compute the moving average of the histogram, and find the minimum of this average. The point of minimum provides us with the threshold for binarizing the image. The binary conversion [2] is such that it inverts the image, i.e. all dark regions including the cavities between fingers get bright whereas the hand region becomes black
6. Application of the morphological operators removes very small connected regions including any holes in the white connected regions as shown in Fig.2.
7. Search for the cavities on the left side of the binary image. Rotating the hand the two cavities– one between the little finger and the ring finger, and another between the middle finger and the index finger are detected.
8. The Laplacian edge detector is applied on the fingers to get the contours of the cavities.
9. A tangent is drawn between the two contours of the cavities in Fig. 4 such that all the points of both the contours lie on one side of the tangent.
10. The perpendicular bisector of this tangent is taken as the x-axis and the tangent is taken as the y-axis to form a coordinate system that facilitates earmarking the ROI.
11. The ROI is rotated and resized to save as a file.



Fig. 1: A typical Palmprint



Fig. 2: Cropped section of image in Fig.1



Fig. 3: Binarized Image

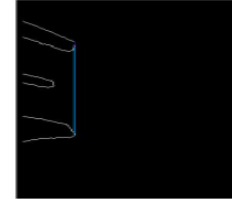


Fig. 4: A common tangent between two curves

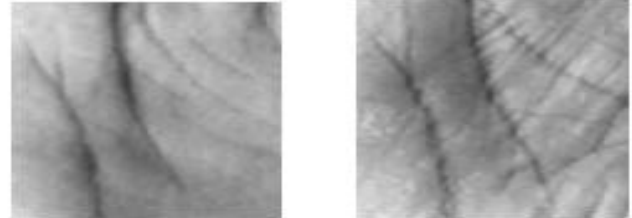


Fig. 5: The extracted ROI of palmprint

## III. FEATURE EXTRACTION

### A: Sigmoid Features

Here the sigmoid function is used to generate features.

1. The average intensity  $I_{avg}$  is calculated from

$$I_{avg} = \frac{\sum_{i=1}^m \sum_{j=1}^n I(i,j)}{m \times n} \quad (1)$$

2. The maximum intensity,  $I_{max}$  is found.
3. The sigmoid feature indicated by sig is obtained as

$$\text{Sig} = \frac{I_{avg}}{1 - \exp\left(-\frac{I_{avg}}{I_{max}}\right)} \quad (2)$$

### B: Energy Features

This feature gives the distribution of energy in the sub windows of ROI. Energy features are derived as follows:

1. The average intensity  $I_{avg}$  is now modified to remove the center pixel intensity

$$I_{avg} = \frac{\sum_{i=1}^m \sum_{j=i}^n I(i,j) - C}{(m \times n) - 1} \quad (3)$$

$C$  is the gray level intensity of center pixel.

2. The maximum intensity,  $I_{max}(i, j)$  is found .
3. The fuzzy membership function  $\mu$  is taken as

$$\mu_0 = 1 - \frac{|I(i,j) - I_{avg}|}{I_{max}} \quad (4)$$

4. The Energy is computed from

$$E_g = \frac{\sum_{i=1}^m \sum_{j=1}^n \mu_{ij}^4}{m \times n} \quad (5)$$

## 3.3:Entropy Features

Entropy is a measure of information of source symbols which is a case of palmprint turn out to be gray levels.

1. The maximum intensity  $I(i, j)$  is show in equation 3
2. The member function is shown in equation 4

## IV. MATCHING

## A: Euclidean Distance Measure

Given two data sets of features corresponding to the training and testing samples, a matching algorithm ascertains the degree of similarity between them. The Euclidean distance is adopted as a measure of dissimilarity for matching the palmprints. Genuine scores are derived from the matching distances (Euclidean distance) between the same user and the imposter scores are calculated from the matching distances between two different users. The graph of receiver operating characteristic (ROC) is a plotted as GAR (Genuine Acceptance Rate) Vs. FAR(False Acceptance Rate)[9,10,11].

## B: Support Vector Machines

SVM operates on the principle of Structural Risk Minimization (SRM) [14, 15]. It constructs a hyper-plane or a set of hyper-planes on a high dimensional space for the classification of input features. Considering a two-class problem to be solved by a SVM, we start with a training sample described by a set of features  $x_i \in R_n$ ;  $n$  being the number of features belonging to one of the two classes designated by the label  $y_i \in \{+1, -1\}$ . The data to be classified by the SVM may not be linearly separable in the original feature space. In the linearly non-separable case the data is projected onto a higher dimensional feature space using Kernel function [8]

$$K(x_i, x_j) = \phi(x_i)^T \cdot \phi(x_j) \quad i, j=1, 2, 3, \dots, N \quad (7)$$

Where  $\phi$  is the function that maps the data on to the higher Dimension space  $H$ . Next, SVM generates a hiper plane  $H$  with the decision boundary given by,

$$f(x) = \sum_{i=1}^N y_i \alpha_i K(x, x_j) + b \quad (8)$$

where  $\alpha_i$  is the non negative Lagrange multiplier. A quadratic optimization problem is set up to solve for the unknown parameter  $s$ . It is defined as:

Minimize:

$$L(w, b, \alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \alpha_i \alpha_j k(x_i, x_j) \quad (9)$$

Subject to:  $0 \leq \alpha_i \leq c$ ,  $i = 1, 2, 3, \dots, N$

Where cost parameter  $c$  control the trade off between the training errors and rigid margins. A few function used in this study are:

Linear Kernel:

$$K(x_i, x_j) = ax_i^T \cdot x_j + b \quad (10)$$

Radial basis function( RBF):

$$K(x_i, x_j) = \exp(-r\|x_i - x_j\|) \quad (11)$$

Sigmoid Kernel:

$$k(x_i, x_j) = \tanh(ax_i^T \cdot x_j + b) \quad (12)$$

Polynomial Kernal function:

$$K(x_i, x_j) = (ax_i^T x_j + c)^d \quad (13)$$

Constants  $a$ ,  $b$ ,  $c$  and the degree of polynomial  $d$ , which is varied in the polynomial kernel function. For the classification of the data LIBSVM [9] has been used. The values of the parameters are selected as  $a=1$ ,  $c=0$ ,  $T=1$  and  $d=1, 2$  and  $3$

## V CONCLUSIONS &amp; FUTURE WORK

A palmprint based biometric authentication system has been developed. Three feature types are experimented on the palmprints for suitability and usefulness for the palmprint based authentication. Of these the sigmoid and entropy features are found to be most suitable in terms of accuracy among all the feature types tested on the two databases. The energy features are lagging behind these two types of features. The recognition rates of 100% are achieved with both sigmoid and entropy features on both the databases using SVM classifier. Irrespective of the type of features used, there is a marked difference in the results obtained using the Euclidean distance classifier and those with SVM classifier. SVM classifier with a linear kernel function and two polynomial kernels of degrees 2 and 3 perform with an accuracy of 100% recognition score on Poly U using sigmoid and entropy features. The main problem for achieving very good authentication rates lies in the choice of the number of samples for both the training and the testing. We have made several experiments by varying these numbers. It is observed that as the training samples increase the matching scores increase but as the number of testing samples increase the matching scores decrease correspondingly. The *cross validation* is also done and the results are almost the same.

The future work will be concerned with developing new classifiers and new features along with the new membership functions.

## REFERENCES

- [1] Xiang-Qian Wu, Kuan-Quan Wang, David Zhang, "Wavelet Based Palmprint Recognition", First international conference on Machine learning and cybernetics, Beijing, pp-1253-1257, 4-5 Nov 2002.
- [2] R. C. Gonzalez, R. E. Woods, Digital Image Processing, Addison Wesley publishers, 1993.

- [3] L. Hong and A. K. Jain, "Integrating Faces and Fingerprints for Personal Identification", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 20, No. 12, pp. 1295-1307, Dec1998.
- [4] Anil K. Jain, Arun Ross and Salil Prabhakar, "An Introduction to Biometric Recognition", IEEE Trans. on Circuits and System for Video Technology, Special Issue on Image and Video- Based Biometrics , Vol. 14, No. 1, January 2004.
- [5] N. Duta, A.K. Jain, and K.V. Mardia, "Matching of Palmprint," Pattern Recognition Letters, Vol.23, No. 4, pp. 477-485, 2001.
- [6] David Zhang, Wai-Kin Kong, Jane You, and Michael Wong, "Online Palmprint Identification", IEEE Trans. Pattern Analysis & Machine Intelligence, Vol. 25, No. 9, pp. 1041 – 1050, September 2003.
- [7] H. K. Polytechnic University, "Palmprint database", Biometric Research Center Website. 2005. <http://www4.comp.polyu.edu.hk/~biometrics/>.
- [8] H. Li, Y. Liang, and Q. Xu, "Support Vector Machines and its Applications in Chemistry", Chemometrics and Intelligent Laboratory Systems, vol. 95, pp. 188-198, Feb. 2009.
- [9] Chih-Chung Chang and Chih-Jen Lin, LIBSVM: a library for support vector machines, 2001. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [10] Li Fang, Maylor K.H. Leung , Tejas Shikhare, Victor Chan, Kean Fatt Choon , "Palmprint Classification", IEEE Int. Conf. on Systems, Man and Cybernetics, Vol. 4, 8-11 October 2006, pp.2965-2969.
- [11] E. Boonchien, W. Boonchieng, and R. Kanjanavani, "Edge-Detection and Segmentation Methods for Two-Dimensional Echocardiograms", Proc. Int'l Conf. Computers in Cardiology, pp. 541-544, September 2004.
- [12] Guang Dai and Changle Zhou, "Face Recognition Using Support Vector Machines with the Robust Feature", in Proc. RO-MAN'03, 2003, pp. 49-53.
- [13] Michael Goh Kah Ong, Connie Tee and Andrew Teoh Beng Jin, "Touch-less Palm-print Biometric System", in Proc. VISIAPP'08, 2008, pp. 423-430.
- [14] V.N. Vapnik, Statistical Learning Theory. New York: Wiley- Interscience, 1998.
- [15] B. Scholkopf, C.J.C. Burges, and A.J. Smola, Advances in kernel Methods-Support Vector Learning. Cambridge, MA: MIT Press, 1998.
- [16] M. Hanmandlu, H.M. Gupta, Neha Mittal, and S. Vasikarla, "An Authentication System Based on Palmprint", in Proc. ITNG, 2009, pp.399-404.

# Image Inpainting-Automtic Detection And Removal Of Text from Images

A. Uday Modha, B. Preeti Dave

**Abstract**—Image Inpainting is the art of filling in missing data in an image. The purpose of inpainting is to reconstruct missing regions in a visually plausible manner so that it seems reasonable to the human eye. There have been several approaches proposed for the same. The novel contribution of the paper is the combination of the inpainting techniques with the techniques of finding text in images and a simple morphological algorithm that links them. This combination results in an automatic system for text removal and image restoration that requires no user interface at all. Examples on real images show very good performance of the proposed system and the importance of the new linking algorithm.

**Index Terms**—inpainting, text detection, mathematical morphology.

## I. INTRODUCTION

Inpainting is the process of filling-in missing or damaged image information. Its applications include the removal of scratches in a photograph [1], repairing damaged areas in ancient paintings, recovering lost blocks caused by wireless image transmission, image zooming and super-resolution [2], removing undesired objects from an image [3], or even perceptual filtering [4]. In the inpainting problem, the user has to select the area to be filled in, the inpainting region, since the areas missing or damaged in an image cannot be easily classified in an objective way. However, there are some occasions where this can be done. One such example, are the text characters printed in an image. Detecting and recognizing these characters can be very important in bimodal search of internet data (image and text), and removing these is important in the context of removing indirect advertisements, and for aesthetic reasons. In this paper we deal with the problem of *automatic* inpainting-based image restoration after text detection and removal from images. That is, given an image with text characters how to detect the exact position of the characters in the text, remove them and fill-in the resulting gaps via inpainting methods. Our contribution is a new system for automatic text detection and removal from images that requires no user interaction. This system is important because the selection of the area to be inpainted has been done manually by previous inpainting systems.

A. Student of Master of Engineering in Information Technology, Shantilal Shah Engineering College, Bhavnagar, (e-mail: udaymodha@gmail.com).

B. Department of Electronics and Communication, Shantilal Shah Engineering College, Bhavnagar, (e-mail: preetidave@yahoo.com).

## II. BACKGROUND AND REVIEW OF INPAINTING

For the inpainting problem it is essential to proceed to the discrimination between the structure and the texture of an image. As *structure* we can define the main parts - objects of an image, whose surface is homogeneous without having any details. As *texture* we can define the details on the surface of the objects which make the images more realistic.

## III. OVERVIEW ON TEXT DETECTION

Much research work has been carried out in the area of text detection and localization in both, images and video. Therefore, for a better understanding of the different methods, the main character features are described. Although hardly ever all of them are taken into account in the same method, some of them, such as contrast or colour homo geneity, are always present. Text extraction from images and video sequences finds many useful applications in document processing [5], detection of vehicle license plate, analysis of technical papers with tables, maps, charts, and electric circuits [6], identification of parts in industrial automation [7], and content- based image/video retrieval from image/video databases [8], [9]. Educational and training video and TV programs such as news contain mixed text-picture-graphics regions. Region classification is helpful in object-based compression, manipulation and accessibility. Also, text regions may carry useful information about the visual content. However due to the variety of fonts, sizes, styles, orientations, alignment effects of uncontrolled illuminations, reflections, shadows, the distortion due to perspective projection as well as the complexity of image background, automatic localizing and extracting text is a challenging problem. Characters in a text are of different shapes and structures. Text extraction may employ binarization [10], [5], [13] or directly process the original image [12], [13], [14]. In [9], a survey of existing techniques for page layout analysis is presented. Mathematical morphology is a topological and geometrical based approach for image analysis. It provides powerful tools for extracting geometrical structures and representing shapes in many applications. Morphological feature extraction techniques have been efficiently applied to character recognition and document analysis, especially if dedicated hardware is used. In this paper, we propose an

algorithm for text extraction based on morphological operations

#### IV SYSTEM FLOW

In this paper, we propose a text detection algorithm that consists of following steps:

- 1) Morphological edge detection
- 2) Text feature filtering
- 3) Text region binarization
- 4) Applying Inpainting Algorithms

##### A. Step 1 - Morphological Edge Detection

To perform the edge detection algorithm, we first convert the input RGB color image to a gray-scale intensity image  $Y$  using (1), where  $R$ ,  $G$ , and  $B$  represent red, green and blue components of the input image.

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

The gray-scale image is then blurred using open-close and close-open filters to reduce false edge noise and over-segmentation. Structuring element used for this operation is a 3 by 3 8-connected element. Next, a morphological gradient detection operation is performed on the blurred image  $Y_{bl}$ , as shown in (2).

$$Y_2 = MG(Y_{bl}, B) = dilation(Y_{bl}) - erosion(Y_{bl}) \quad (2)$$

In order to get the threshold level of  $Y_2$ , we use a global nonhistogram-based thresholding technique. The threshold level is determined by (3), where  $s$  is an edge detector obtained by applying central difference edge detection filter to  $Y_2$ . [1]

$$\gamma = \frac{\sum Y_2 \cdot s}{\sum s}$$

##### B. Step 2 - Text Feature Filtering

In order to reduce the number of connected components that have to be analyzed, a close operation with a 5 by 5 structuring element is performed to the binary edge image obtained from Step 1. After the close operation, all connected components of the edge image are Screened with their position, size, and area information. A candidate of letter should meet a set of constraints in size and shape. In our algorithm, we select connected components as letter candidates if the following requirements are met:

- 1) Width of the bounding box  $< 0.5$  image width
- 2) Height of the bounding box  $< 0.3$  image height
- 3)  $0.1 < \text{center width of the bounding box} < 0.9$
- 4)  $0.3 < \text{center height of the bounding box} < 0.7$

- 5) Width vs. height ratio  $< 10$
- 6) Width of the bounding box  $> 10$  pixels
- 7)  $0.1 < \text{Connected component filled area over (width height of the bounding box)} < 0.95$
- 8) Width of the bounding box  $> 10$  pixels
- 9)  $0.1 < \text{Connected component filled area over (width height of the bounding box)} < 0.95$

After the first round filtering, it is expected that most of the non-letter components would be removed. So the majority of the remaining candidates should be letters with the same font and size. Based on this condition, we calculate the mean height  $h_m$  of the bounding box of the remaining components, and remove any connected component with its height smaller than  $0.6h_m$  or greater than  $1.8h_m$ .

##### C. Step 3 - Text Region Binarization

Each remaining bounding box is used as a mask to the original gray-scale image. Otsu's method [15] is used to obtain the threshold of the masked gray-scale image for binarization. Since each bounding box is relatively small compared to the size of the entire image, no further adaptive thresholding method is implemented. Theoretically after this step, only stroked letters are left as the foreground, 1, and the rest of the image would go to background, 0.

##### D. Step 4 – Applying Inpainting Algorithms

As a Final Step of our system, We now use these positions derived above as the inpainting area and apply any of the inpainting technic to remove the text from images. Result of step 3 is a binary image. Which works as a mask in inpainting technic.

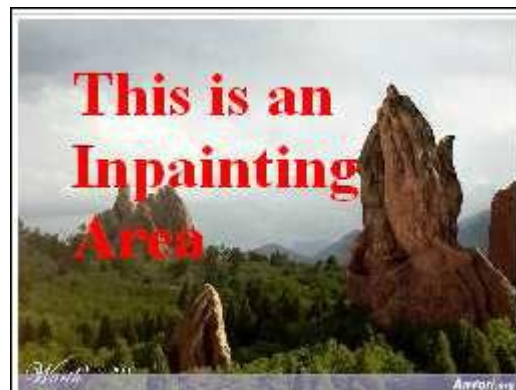


Figure 1 Original Image





Figure 2 Text Detection

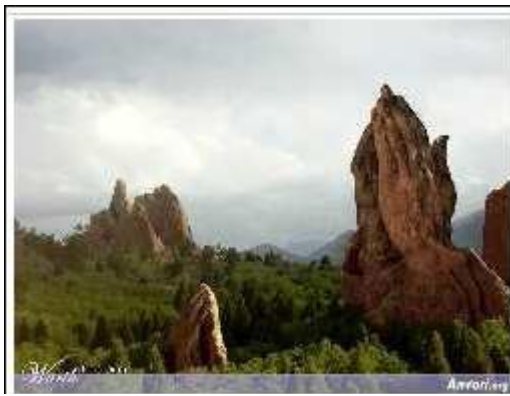


Figure 3 Inpainted Image

## V. EXPERIMENTAL RESULTS

We present results of the system that we described above. It is important to note that, the system requires no user interaction, since both the text detection (inpainting region) and the inpainting process are being done automatically. Fig. 1 is the Original Image which is to be inpainted. In Fig. 2 we used Morphological Text Extraction algorithm from [16] to detect the text position in the images. And works as a mask in inpainting algorithm. In Fig. 3 is the Final inpainted image after application of appropriate inpainting algorithms. In general, an automatic system is needed to decide about the appropriate inpainting method.

## VI. CONCLUSIONS

In this paper we dealt with the combined problems of inpainting and finding text in images. We have proposed a new simple morphological algorithm inspired from the reconstruction opening operation. This algorithm captures all the pixels that correspond to text characters and thus its output can be considered as the inpainting region. By applying then an appropriate inpainting method, we have

developed a system for automatic text removal and image inpainting.

## REFERENCES

- [1] M. Bertalmio, V. Casseles, C. Ballester, and G. Sapiro, *Image Inpainting*, ACM Press, 2000.
- [2] T. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM J. on Appl. Math.*, vol. 62, pp. 1019–1043, 2002.
- [3] A. Criminisi, P. Perez, and K. Toyama, "Object removal by exemplar based inpainting," in *Proc. IEEE-CVPR*, 2003, vol. 2, pp. 721–728.
- [4] M. Dimmiccoli and P. Salembier, "Perceptual filtering with connected operators and image inpainting," in *Proc. ISMM*, 2007.
- [5] K. Jain and B. Yu, "Document representation and its application to page decomposition," *IEEE Trans. Pattern Anal Machine Intell.*, vol. 20, pp. 294–308, Mar. 1998.
- [6] K. Jain, and B. Yu, "Automatic Text Location in Images and Video Frames", *Pattern Recognition*, 31 (12) (1998) 2055-2076.
- [7] K. Jain, and Y. Zhong, "Page Segmentation using Texture Analysis", *Pattern Recognition*, 29 (5) (1996) 743-770.
- [8] C. Gopalan and D. Manjula, "Contourlet Based Approach for Text Identification and Extraction from Heterogeneous Textual Images", *International Journal of Computer Science and Engineering* 2(4) (2008) pp.202-211.
- [9] D. Crandall, S. Antani, and R. Kasturi, "Robust Detection of Stylized Text Events in Digital Video", *Proceedings of International Conference on Document Analysis and Recognition*, 2001, pp. 865-869.
- [10] K. C. Fan, L. S. Wang, and Y. K. Wang, "Page segmentation and identification for intelligent signal processing," *Signal Process.*, vol. 45, pp. 329–346, 1995.
- [11] H. M. Suen and J. F. Wang, "Text string extraction from images of color-oriented documents," *Proc. Inst. Elect. Eng. Vis., Image, Signal Process.*, vol. 143, no. 4, pp. 210–216, 1996.
- [12] R. Lienhart, "Indexing and retrieval of digital video sequences based on automatic text recognition," in *Proc. ACM Int.*
- [13] L. Wang and T. Pavlidis, "Direct gray-scale extraction of features for character recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1053–1067, Oct. 1993.
- [14] Y. Zhong, K. Karu, and A. K. Jain, "Locating text in complex color images," *Pattern Recognit.*, vol. 28, no. 10, pp. 1523–1535, 1995.
- [15] N. Otsu, "A threshold selection method from gray-level histograms". *IEEE Trans. Sys., Man., Cyber* 9(1):62-66
- [16] Y. M. Y. Hasan and L. J. Karam 2000 "Morphological Text Extraction from Images" *IEEE Transaction on Image Processing* vol. 9, no. 11, Nov. 2000.
- [17] M. M. Oliveira, B. Bowen, R. McKenna and Y. C. Chang, "Fast Digital Image Inpainting", in *Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP 2001)*, Marbella, Spain, September 3-5, 2001
- [18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Pearson Education, 2002.
- [19] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald, P. J. W. Rayner "Interpolation of missing data in image sequences", in *IEEE Transactions on Image Processing* 11(4), pp. 1509-1519, 1995.

# A Survey of Techniques in MPLS Networks

A. Ila Vaghela, B. Rahul A Vaghela

**Abstract:** *Multi-protocol Label Switching (MPLS) has become an attractive technology of choice for Internet backbone service providers. MPLS features the ability to perform traffic engineering and provides support for Quality of Service traffic provisioning. This has resulted in increasing interest in MPLS network reliability and survivability. This paper will survey the various techniques utilized in MPLS.*

**Index Terms**—Multiprotocol Label Switching, Virtual Private Network

## I. INTRODUCTION

AS demand grows for the Internet to bring more traffic and provide support for Quality of Services. It is vital to maintain a high level of performance and efficiency. Traffic engineering is the process of optimization of the network to maximize performance and efficiency. Multi-Protocol Label Switching (MPLS) is a tool for network traffic engineering and hence is becoming the technology of choice for Internet backbone.

Network resources also vary due to new resource requests or topology changes. One important part of designing a quality of service network is the reliability of the network. It could be provided with different fault management mechanisms applied at different network levels and timescales. Multi-protocol label switching (MPLS) provides a fast restoration method to recover from failures. MPLS fault restoration mechanisms use backup label switch path (LSP) establishment. With these backups, traffic could always be redirected when a failure occurs. MPLS also provides fault detection and fault recovery actuation faster and more efficiently than other network protocols or technologies.

MPLS forward packets based on fixed size labels, the path that the packets traverse is pre-established. Incoming packet labels are examined to determine the next hop; the old label is then exchanged with a new label to be forwarded to the next hop. The path the packet traverses is called a Label Switch Path (LSP) and the Label Distribution Protocol (LDP) is used to set up the LSP. LSP can allow packets to traverse an explicit route compare with dynamic routing in IP, allowing for traffic engineering. Another key feature of MPLS is the functionality of Label Stacking, allowing a packet to have multiple labels at one time. Therefore packets can be tunneled through several LSPs.

The reservation of network resources is also possible using

MPLS, thus providing QoS guarantees. Two signaling protocols have been established to provide the reservation of bandwidth for the LSP, resource reservation protocol with traffic engineering extension (RSVP-TE) and constraint based- routing label distribution protocol (CR-LDP).

This paper will discuss all the techniques utilized to provide more quality of service using MPLs. It will also survey the different protection and recovery techniques utilized.

## II. 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. [3] 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
- ✓ Multiprotocol Support

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

- ✓ Guarantee a fixed amount of bandwidth for various applications
- ✓ Control latency.
- ✓ Provide quantifiable SLA.
- ✓ 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-

A. Lecturer, GIT, Gandhinagar (ila.vaghela@git.org.in)

B. HoD, IT Department, GIT, Gandhinagar (rahul.vaghela@git.org.in)

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.

### III. MPLS Maneuver

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.

#### A. Maneuver

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.

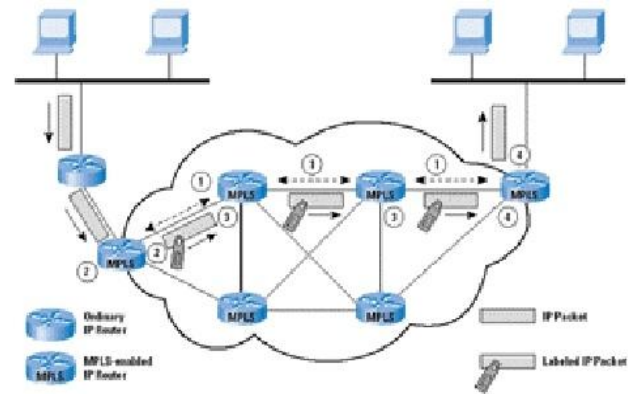


Fig.1 MPLS Operations [1]

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 code points 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 sent 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]

#### *B MPLS Protection Methods*

Protection methods follow a cycle, starting when the fault is detected and finishing when the LSP is recovered. This cycle involves the development of two main components: a method for selecting the working and protection paths, and a method for bandwidth reservation in these paths. A fault detection mechanism along a path and a fault notification mechanism are also necessary to convey information to the network entity responsible for reacting to the fault and taking appropriate corrective actions. Finally, a switchover mechanism to move traffic over from the working path to the protection path can also be provided.

The usual method of offering protection in MPLS environments is to pre-establish a backup LSP onto which to switch traffic when failure occurs. Backup LSP types may be different depending on where they have originated or what types of failure/recovery notification are Global repair Model and Local segment Restoration (Local repair).

#### *C Routing Information*

The basic information needed by any routing protocol to make appropriate path selection decisions is the state of the network. Every routing protocol uses this information to forward packets. The information about the state of the network includes the network topology along with resource availability for QoS purposes. Each change in the state of the network should be detected and disseminated to all the routers in the same autonomous system (AS) and also propagated across AS boundaries until all ASs have been informed of this

change. The main cause for state change is resource availability variation in the network since topology variations are less frequent. The large amount of information exchange for state update can compromise the scalability of the routing schemes. To reduce this amount, two approaches are possible: reducing either the frequency of updates or the details in the updates. The former is achieved by using various mechanisms such as class-based, threshold-based, and periodic updates. The latter is achieved by aggregating the network state information. In the case of MPLS networks, a centralized network manager can also be used for network operation, in which case the problem of information dissemination becomes redundant.

#### *D MPLS Recovery Mechanism*

"To deliver reliable service, MPLS requires a set of procedures to provide protection of the traffic carried on different paths. This requires that the label switching routers (LSRs) support fault detection, fault notification, and fault recovery mechanisms, and that MPLS signaling support the configuration of recovery"[1]. As mentioned earlier there are two main signaling protocols for MPLS networks which support the configuration of recovery. RSVP-TE is based on the extension of the reservation protocol to support traffic engineering. This is a soft state signaling protocol, requiring periodic refresh messages to ensure links are still active. The CR-LDP is an extension based on the new Label Distribution Protocol (LDP). CR-LDP also allows for reservation of bandwidth and is a hard state signaling protocol, ie does not require refresh messages to maintain an active link. Chung provides a comparison of the two signaling protocols [2]. Fault detection is the time required to detect a fault; the quicker the fault is detected in the network the earlier the fault can be repaired. The current method of detecting faults on a link comes from the use of periodic messages, if the messages are absent then the link is assumed to be faulty and fault notification is initiated. Fault notification is the process of informing other LSRs of the failure that occur which eventually propagates to ingress LSRs to initiate recovery procedures to maintain traffic flows.

The recovery models used in an MPLS network are classified into two categories, rerouting and protection switching. The rerouting recovery model can only establish the recovery path when a fault has occurred. The new recovery path is usually created based on fault information, network topology, etc. Traffic is switched onto the recovery path temporarily until the fault is repaired. Rerouting has the advantage of using up-to-date information and is able to make efficient use of resources in creating recovery paths. The drawback is the time required to compute and setup the recovery path after failure has occurred. Figure 1 shows an example of link failure between node C and D. The rerouting recovery model sends a Fault Indicator Signal (FIS) upstream to ingress node A, which then finds and establishes a recovery path around the failed link, eg. From A → F → G → H. [2]

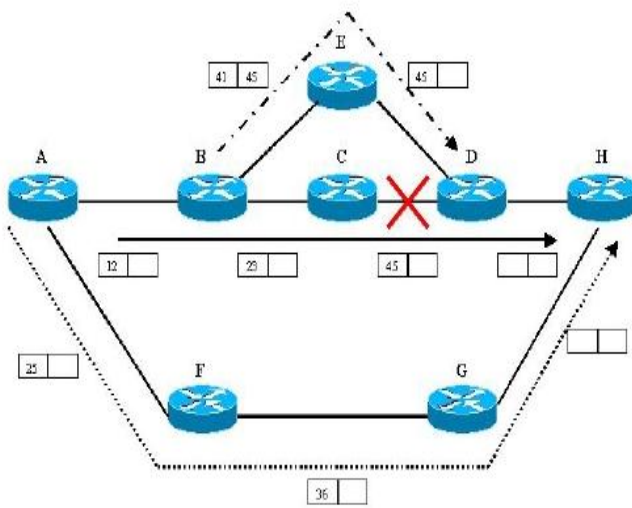


Fig. 2 Example of MPLS Recovery [2]

Protection switching involves pre-determining and establishing a recovery path before the fault has occurred. The recovery path is usually disjoint from the working path to prevent any fault that can affect both the working and recovery path; When failure occurs the traffic is switched from the working path to the recovery path, this is simply done by swapping the label from the working LSP with the label associated with the recovery LSP. This has the advantage of permitting fast recovery, but requires additional resources to setup the recovery path which remain unused until a fault occurs. Protection switching can also be divided into local or end-to-end repair. Local repair is used to protect against the fault of a single link or node, it has the advantage of minimizing the time required for fault notification, as only the previous upstream node is notified of the fault. The upstream node creates a recovery path around the fault and label stacking is used to direct traffic onto the recovery path. Once traffic has been routed around the fault the top label is removed and traffic continues as normal without the ingress or egress LSRs ever knowing a fault has occurred. An end-to-end repair requires the recovery path to protect against all links and nodes used in the working path. An example of local and end-to-end repair is shown in Figure 1. If link failure occur between node C and D, protection switching using end-to-end repair would established a recovery path  $A \rightarrow F \rightarrow G \rightarrow H$  and reserve the necessary resources. Local repair shows that the upstream node B established a recovery path  $B \rightarrow E \rightarrow D$  to protect node C. Label stacking is used to divert traffic to the recovery path by simply adding an extra label to the packet at node B. Label 41 is added at node B hence traffic is switched to the recovery path, at node E the extra label is removed and traffic is switched back to the original path. [2]

#### IV CONCLUSION

From the study of various operation, protection and recovery techniques of Multi protocol Label Switching we can conclude that it provides very high level functionality to support all the requirements of the Service Provider. It discusses the global and local repair model for better protection techniques and rerouting and protection switching techniques for the restoration. Altogether, it is a very innovative methodology in combination with layer 2 and 3 technologies and in foundation having the concepts of switching and routing technology at rim. It provides enhanced routing and forwarding scalability and traffic engineering capabilities for better network provisioning.

#### REFERENCES

- [1] J. Sharma , F. Hellstrand,, "Framework for multi-protocol label switching (mpls) based recovery," 2003, request for comments: 3469.
- [2] J. Foo,"A Survey of Service Restoration Techniques in MPLS Networks",
- [3] F. Le Facheur , "Requirements for Support of Diff- Serv-Aware MPLS Traffic Engineering," IETF RFC 3270, May 2002.
- [4] W. Huang., J. Yang.,," New VPN Application in 3G Network", ISBN 978-952-5726-07-7,2009
- [5] J. Marzo, E. Calle, A. Tricha , "QoS Online Routing and MPLS Multilevel Protection: A Survey", 2003
- [6] E. Rosen, R. Viswanathan , "A Proposed Architecture for MPLS," draft-ietf-mpls-arch-01.txt, March 1998.
- [7] T. Kolar, A. Valencia., M. Littlewood, "Layer Two Forwarding (Protocol) L2F," draft-valencia-l2f-00.txt, , October 1997.
- [8] J. Heinanen, E. Rosen. "VPN Support for MPLS," draft-heinanen-mpls-vpn-01.txt, March 1998.
- [9] P. Ferguson, G. Huston, "What is a VPN", April,1998

# Usability Testing on Prototype Design of University Blackboard Tool: Effective Channel of Communication and Collaboration

A. Birendrasinh K. Zala

**Abstract**—A major challenge, and thus opportunity, in the field of human-computer interaction and specifically usability engineering is designing effective user interfaces for emerging technologies that have no established design guidelines or interaction metaphors or introduce completely new ways for users to perceive and interact with technology and the world around them. We present an exemplar user-based study conducted on prototype design of University Blackboard Tool which is developed to establish effective channel of communication and collaboration. This paper will discuss all steps of usability testing in a very general manner. The research method used in this method is Formative Evaluation of prototype.

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

## I Introduction

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. UCD approach includes System Description activity, System Redesign activity and Evaluation activity.

A. AP, IT Deptt., Gandhinagar Institute of Technology, Gandhinagar, (e-mail: birendra.zala@git.org.in).

- 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.

This paper discusses steps which performed during the formative evaluation of University Blackboard Tool. Formative Evaluation (both formal and informal) is an observational, empirical evaluation method that assesses user interaction by iteratively placing representative users in task-based scenarios in order to identify usability problems, as well as to assess the design's ability to support user exploration, learning, and task performance[4]. This paper also examines the results of evaluation phase

## II Case Study: A User-Based Study Examining University Blackboard Tool

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.

### A. System description phase

The Group Work Tool (Blackboard Add-on) shall support students, lectures and tutors for academic purposes where group or team work is involved. This is a type of groupware tailored suit the existing academic environment. Blackboard tool relies on the Internet. Also group related data can only be accessed by relevant group members. However, in some subjects the convener or tutor needs to assess each group member's contribution by viewing the group area. Assigned staff members for the subject should be able to carry out this assessment as well.

This groupware should be able to provide non-located users to interact in synchronous or asynchronous manner. This product has to cater groups of varied interaction styles as group composition changes from one to another and it changes from assessment item to another. Therefore different communication mechanisms need to be present in this tool.

- Share messages among members, replying messages.
- Email facilities
- Instant messaging.

Moreover product shall support sharing documents via a central repository of documents and versioning of progressive documents. This product shall support setting appointments, managing events and calendar capabilities to setup meeting times so that students can perform their task in a timely and informed manner. In addition it allows task allocation for individual students so that group as a whole knows what the responsibilities of each student students

**B. System Redesign Phase**

System redesign ensures that an interactive design [5] 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:

- Table.1 shows essential use case for justification of how the different tasks have been allocated between the user and the system. Example activity mention here is “chat with group members”.
- Table.2 details a concrete use case corresponding to the essential use case. Mark up the concrete use case, indicating the task objects and attributes.
- Prototype will be produced as a main deliverable and will serve as a foundation for Evaluation phase. Figure 1 shows functional paper prototype [6] design for the activity of “chat with group members”.

Table 1. Essential use case – Chat with group members

User’s Purpose	System Responsibility
Identify himself	Prompt for login information.
	Verify student identity.
	Display university student home
View subject	Display subject content.
View group details	Display group related information.
Start chat	Display an area for chatting.
	Display group members.
Identify the member to chat	Indicate which member is being talked to.
	Display an area to enter the message
Send message	Display what is typed in.
	Display progress of sending.
	Display the other group member has

Table 2. Concrete use case - Chat with group member

User’s Action	System Responsibility
<u>Student</u> opens a browser and opens the university home page and clicks on blackboard	System requests for <u>student id</u> and <u>password</u>
<u>Student</u> enters their <u>student id</u> and <u>password</u>	System requests for <u>student id</u> and <u>password</u>
	System verifies the identity.
<u>Student</u> selects desired <u>subject</u>	System displays a list of options a student can choose from.
Student selects <u>group</u> option	System displays a list of features in group area and Student names.
Student selects to chat	System display an area for chatting
	System displays the list of members in the group
Student select a <u>member</u> to chat to	System highlights the <u>member</u> in the list.
Student types a message and sends	System displays what is typed in.
	Display progress of sending.
	Display the other group <u>member</u> has received the <u>message</u> . With details such as <u>time</u> , <u>date</u> , <u>content</u> and <u>sender</u>

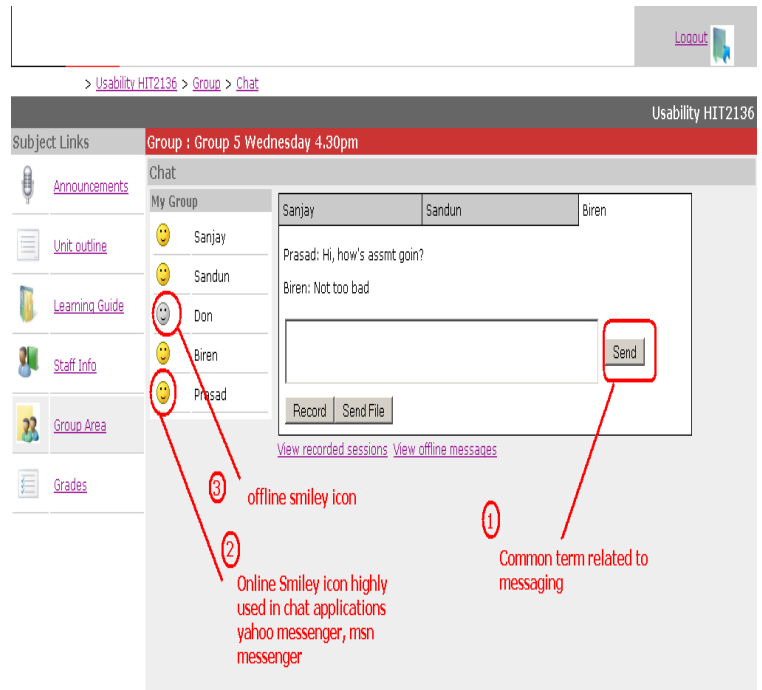


Figure.1. Paper Prototype Design for task “chat with group members via Blackboard tool”

Metaphors used (by numbers, fig 1)

- 1 Send

Common chat applications use this term to indicate sending of a message.

2, 3 Online smiley icon / offline smiley icon

Popular in most web applications, used as indication of status to be online/offline in chat applications like yahoo. They can be easily recognizable.

2 Record

This term indicates persistence, saving something so that it can be retrieved later. This simple term is used on many applications like radio, TV.

**C. SYSTEM EVALUATION PHASE**

In usability evaluation phase followed activities conducted on prototype design of “Chat with group members “ The term usability evaluation is here used to describe the complete test of the UI, including planning the evaluation, conducting the evaluation sessions and presenting the results.

Task description was created as shown in figure 2 for use in usability test. Justify the task/s to be carried out.

**Task Description**

In this task we would like you to chat with a member of your group member. You are in the chat page in your subject's group area. You would start by opening chat.htm in a web browser.

Information you may need to complete this task:

We like you to send following messages to your selected group member.

Message 1: Hi Birendra, how are you?  
 Message 2: that’s good How’s the assignment?  
 Message 3: Excellent, you need any help? "

When you finish the task please leave the page chat page open

Figure.2. Task Description task “chat with group members via Blackboard tool”

- ➔ Preparation of demographic questionnaire and any extra questions plan to use in the evaluation.
- ➔ Preparation of a copy of observation sheet.
- ➔ Gather and analysis of data gathered during testing.
- ➔ Calculate metrics to identify efficiency and effectiveness of UI.

**III RESULTS**

The table 3 summarizes the data for task 2. All the participants complete the task 2 “Chat with the group member” without any assistance giving a task completion rate of 100%. No errors were found in order to complete the Task 2. The shortest task completion time was 1.08 minutes and the longest was 1.63. The average Task Completion Time was 1.33 minutes.

Table 3: Task Metrics for “chat with group members via Blackboard tool”

Sr no.	Unassisted Task Completion	Errors	Task Completion Time(mins)
P1	100%	0	1.12
P2	100%	0	1.63
P3	100%	0	1.08
P4	100%	0	1.5

**IV Conclusion**

The Usability test was conducted on paper prototype design of University Blackboard tool. As I was closely watching every move of the participants, I observed that finding the group area for the participant was a bit difficult. The link was labeled as “Group Area”. In the task description there was no mention of a group area as such, rather used “group” to identify the group work tool. This may have confused the participant (Fig 1). However, this is what happens in the existing application.

**REFERENCES**

- [1] A. 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,” Behaviour and Information 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] D. Bowman, J. Gabbard, and D. Hix, “A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods,” Presence: Teleoperators and Virtual Environments, vol. 11, no. 4, 2002, pp. 404-424
- [6] E. Ojakaar, (2001, Oct 25),” Users Decide First; Move Second,” UIEtips. Retrieved February 17, 2012. from [http://www.ue.com/articles/users\\_decide\\_first/](http://www.ue.com/articles/users_decide_first/)
- [7] J. M. Spool, (2007, Feb 27), “Usability Testing: Oh, The Things You Can Learn,” UIEtips. Retrieved March 16, 2011.
- [8] [http://www.ue.com/articles/usability\\_tests\\_learn/](http://www.ue.com/articles/usability_tests_learn/)



# A Comparative Analysis of Cancer Gene Classification Approaches from Microarray Data

A. Mahesh Panchal, B. Dr. Sanjay Garg

**Abstract**—Cancer is one of the dangerous diseases and requires long treatment too. In order to detect it in early stage, human genes play an important role. Classification, which is one of the knowledge discovery technique of data mining can able to separate out genes which are responsible for cancer from given microarray database. There are many classification methods and feature selection methods applied before classification. In this paper a survey of those methods is described and applied on given dataset which helps the researches to make comparison among them and some new idea may be clicked.

**Index Terms**— Bioinformatics, Classification, Data Mining, Feature selection, Genes.

## I. INTRODUCTION

Bioinformatics is integrated field whose roots fall mainly into biology, data mining & statistics. Since last few years huge amount of biological data has been generated like DNA database, protein database, gene expression database etc. Biologists are interested in analyzing these data to identify genes for particular disease, to compare the DNA or protein sequence, to predict protein structure. To identify genes which are responsible for cancer disease and to classify genes for different types of cancer is one of challenging tasks due to availability of large number of genes compared to number of samples and presence of noise in gene expression data<sup>[1]</sup>. Before any gene expression database is used for classification, the most relevant features (genes) considering class label attribute (either cancer /normal or type of cancer) must be found using some feature selection method or dimensionality reduction method. This will certainly reduce noisy features from dataset and keep computation and testing burden low<sup>[2][3]</sup>. Using these selected genes, any well known classifier is to be trained to derive the classification model. This model is to be tested on same kind of data but should be different from one which was used for deriving the classification model to ignore the overfitting. If accuracy of model is acceptable then later it can be used for classification of unseen data. Because only highly related genes are used during classification, accuracy of classifier is increased and computational cost is reduced. In this paper, some known feature selection methods & classifiers are applied on given gene expression database and results are analyzed.

A. Author is with Department of Computer Engineering, Kalol Institute of Technology & Research Centre, Gujarat, India (e-mail: mkhpanchal@gmail.com).

B. Author is with Department of Computer Engineering, Nirma University, Gujarat, India.

## II. LITERATURE REVIEW

The problem of identifying most relevant genes for cancer classification has been widely experimented and analyzed. In [4] Yeast gene expression dataset is used which has 2467 genes & 79 experiments performed on them. To select the best features which are relevant to class, some feature selection methods like information gain and dimensionality reduction method like PCA is be used. Three learning algorithms are used: KNN, naïve bayes and decision tree. All three are trained with all 79 features, as well as after applying feature selection method like IG & dimensionality reduction method like PCA. Performance of classifier in each case is measured with three fold cross validation. The best classification accuracy is achieved by IG feature selection of genes with use of KNN and naïve bayes. In [5], classifier is built on ovarian, lung and prostate types of cancer datasets. Whole database is vertically partitioned such that each partition has 2000 genes. On every partition GA-CFS method is applied to select most significant features (genes). Classifiers (DT, SVM, Bagging, stacking) are trained from reduced gene database. It is validated by 10-fold cross validation method. It is shown that GA-CFS increase the accuracy of classifier in each dataset. In [6] DNA methylation array & Affymetrix oligonucleotide array (GED) datasets are used for classification of cancer and normal genes as well as genes for types of cancer. Genetic Algorithms (GA) and Tabu Search (TS) are used for feature selection. SVM classifier is trained with 10-fold cross validation & all 10-folds are constructed randomly 30 times. Three conclusions have been carried out. (i) Gene subset selection achieves higher accuracy than Individual Feature Ranking (IFR) for binary classification. For multi class classification, TS/SVM outperforms IFR. For GED, GA/SVM & TS/SVM outperforms IFR. (ii) Wrappers (Features are evaluated by accuracy of classifier's prediction) are better than filters (MRMR, distance, correlation, consistency). (iii) TS outperform GA in all way.

## III. MICROARRAY TECHNOLOGY AND GENE EXPRESSION DATABASE

Every living organism has DNA in each cell nucleus. It is made up of molecules known as nucleotides. DNA is usually found as two strands that wrap around each other in a spiral within a spiral. A gene is a specific subset of nucleotides on DNA<sup>[7]</sup>. They contain the encoded information to make functional molecules known as proteins. To make proteins from genes is a two step process: (i) transcription (ii) translation. Together it is also known as gene expression. In

first step genes are transcribed into another form known as RNA. Those RNAs are translated into proteins composed of various amino acids. Every biological change come in living organisms are due to alteration inside genes. The genes which are expressed at some level, same may not be expressed at same level under some other environmental condition for example. Microarray is a technique by which expression level of tens of thousands of genes can be made available in a single time <sup>[7]</sup>. After microarray is applied, a database known as gene expression database is created which contains the expression level of thousands of genes. Previously tumor diagnostic methods based on only morphological appearance of the tumor while gene expression database is more accurate and reliable to detect the tumor.

#### IV. FEATURE SELECTION

Real world gene expression database obtained from microarray technology is very much high dimensional. If any classifier is directly trained from all such genes, there is a high possibility of noise, less effectiveness, less accuracy and high computational cost. So, selection of highly informative genes which can best classify given sample is very important <sup>[8]</sup>. There are various advantages of feature selection. First, it will not consider the features which may create noise during classification. Second, it removes redundancy among features and thus less computation burden while training the classification model and testing the unseen data. By feature selection, lots of candidate feature subsets are examined from whole feature space and a subset which is most related to class label attribute can be chosen for classification. There are many feature selection methods available. Each has their own advantages and limitations. All such methods mainly fall under four categories as described below <sup>[6]</sup>.

1. Individual Feature Ranking (IFR): All features (genes for example) are ranked based on their relevance with class label attribute. Most relevant gene is given rank of 1 and so on. This method is simple and scalable too. But it cannot detect correlated genes so redundancy still preserved. Using any of the IFR method, subsets are made of topmost 10, 20 or even 100 genes according to their rank. Then any classifier is trained from all such subsets. Accuracy is measured with n-fold cross validation. A subset achieving highest accuracy is chosen as best feature subset. Some of the examples in this category are information gain, T-score, chi-square, relief-F, symmetric uncertainty.

2. Exhaustive Search: It is performed in whole feature space to generate all possible feature subsets. As this method is generating all possible subsets, it is NP-hard and becomes computationally intractable as number of features increasing.

3. Heuristic (Deterministic): It is a greedy approach which selects or removes features step by step based on local changes and not preferred as it finds local optimum solution. Some of the examples in this category are sequential forward selection, sequential backward selection, best first search.

4. Heuristic (Nondeterministic): It finds optimal or suboptimal solution considering global feature space. It generates subsets in random order contrasting to deterministic approach which generates subsets in step by step manner. Some of the examples are simulated annealing, genetic algorithm.

In Exhaustive Search, Heuristic (Deterministic) and Heuristic (Nondeterministic) methods mentioned above, every time each subset is evaluated by either some filters (distance, consistency & correlation) or by wrappers (any classification algorithm is applied on feature subsets and evaluation is done by measuring accuracy each time). The subset which outperforms others will be given as output.

#### V. INPUT PARAMETERS OF WORK

(I) A publicly available database DLBCL (DIFFUSE LARGE-B-CELL LYMPHOMA) is used here for analysis. Biopsy samples of diffuse large-B-cell lymphoma from 80 patients were examined for gene expression with the use of DNA microarrays and analyzed for genomic abnormalities. This dataset is about prediction of survival after chemotherapy. It contains 80 samples & 7399 genes.

(II) Weka 3.6 (A Practical Machine Learning Tool) is used for all experiments in this paper. It has implemented number of algorithms for classification, attribute selection, discretization, association etc. in Java. One can run any algorithm after loading dataset either in GUI or using command line interface.

(III) Four types of feature selections methods are used to select most relevant genes from database.

1. Information Gain method based on Individual Feature Ranking (IFR)
2. Exhaustive Search gene subset selection
3. Best first search based on heuristic (Deterministic).
4. Genetic Algorithm based on heuristic (Non-deterministic).

In weka, first a feature selection method is chosen among all available like Best First, Exhaustive Search, Ranker etc. After that a evaluation method like CfsSubsetEval, ClassifierSubsetEval, InfoGainAttributeEval etc. is selected which evaluates set of the attributes resulted by feature selection method.

(IV) Two learning algorithms are used to build classification model for best gene subset selected from each of the above feature selection method.

1. Decision Tree
2. Naïve Bayes.

Each classifier's accuracy is measured with 10-fold cross validation. In Weka, there are various categories of learning algorithms like tree, rule, bayes, functions, lazy, meta etc. In tree based algorithms, one is J48 which is java implementation of C4.5. It is enhancement of original ID3 which is based on decision tree. In C4.5 numerical attributes are handled along with missing values in data sets. In any decision tree the classification model is represented by rules. Naïve bayes differs from decision tree in a way that it shows class probability along with rules.

VI. RESULTS AND DISCUSSION

Fig.1 plots the accuracy of classification model derived by decision tree and naïve bayes learning algorithms. Both these algorithms take most relevant genes as input which were generated by Ranking feature selection using Information Gain evaluation method by varying number of genes to generate as output. It can be easily seen that decision tree gives highest accuracy (71.25%) when no. of most relevant genes are 20 and for naïve bayes it is 93.75% for 50 genes. As no. of genes increases accuracy of classifier is almost decreases. Most importantly, naïve bayes always gives higher accuracy than decision tree for given same set of genes.

In Fig. 2, accuracy come from model derived by same two learning algorithms as in Figure 1 is plotted but here difference is to find most relevant genes, genetic search method is used. There is no fix pattern available in chart but it can be infer that by proper selection of population size and no. of generations, good set of attributes come out. In this experiment, cross over and mutation probability is set 0.6 and 0.033 respectively.

Fig. 3 compares the three attribute selection methods for accuracy of two learning algorithms. When no attribute selection method is used then accuracy of both is least. In ascending order according to accuracy given by both classifiers, three methods of attribute selection methods are genetic search, best first search and ranker. Ranker not only generates best attribute set but size of that best set is minimum too (20 for DT and 50 for NB). In genetic search, as shown in Figure 2, to get higher accuracy of 56.25%, DT requires 1920 genes, similar scenario is in NB (2873 genes for 56.25% accuracy). Another point is to be noted here is NB always outperform DT irrespective of any feature selection method.

VII. CONCLUSION

Because of high dimensionality of cancer GED, some feature selection method should be applied to reduce testing overhead and to increase overall accuracy. In this paper a ranker based and two subset methods are applied on a dataset which has more than 7,000 features. On reduced feature sets given by these methods, a classification model is derived using DT and NB. Result shows that ranker based methods are good choice compared to subset based. Moreover, In general NB learning algorithm derives good classification model from reduced set of features. Of course in genetic search by varying no. of generations and population size, more experiments can be carried out.

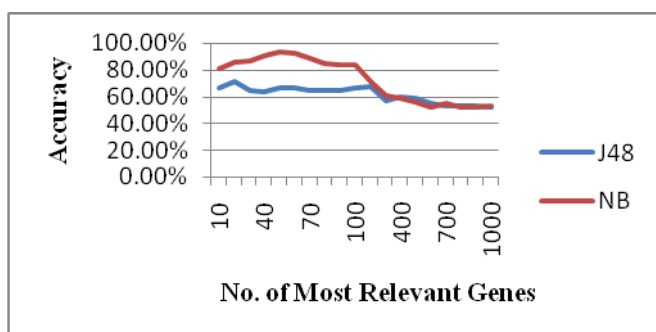


Fig. 1: No. of Most Relevant Genes (Using Ranker-Information Gain) Vs. Accuracy (DT and NB)

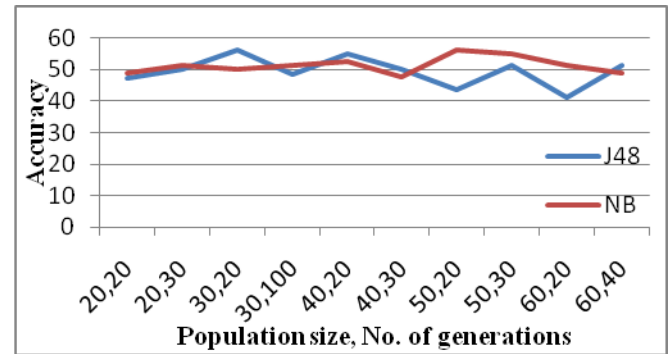


Fig. 2: Population size, No. of Generations (Using Genetic Search-Subset Evaluation) Vs. Accuracy (DT and NB)

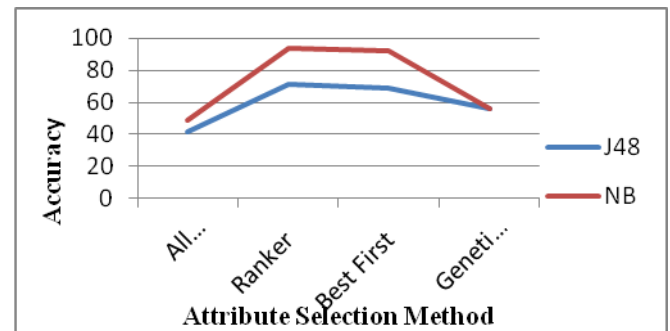


Fig. 3: Attribute Selection Methods Vs. Accuracy (DT and NB)

REFERENCES

- [1] Jinn-Yi Yeh, Tai-Shi Wu, Min-Che Wu, Der-Ming Chang, "Applying Data Mining Techniques for Cancer Classification from Gene Expression Data", *International Conference on Convergence Information Technology*, 0-7695-3038-9/07, IEEE
- [2] Lipo Wang, Feng Chu, and Wei Xie, "Accurate Cancer Classification Using Expressions of Very Few Genes", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol:4 No:1, January-March 2007.
- [3] Xiaosheng Wang1 and Osamu Gotoh, "A Robust Gene Selection Method for Microarray-based Cancer Classification", *Cancer Informatics* 2010:9 15–30
- [4] Changjing Shang and Qiang Shen, "Aiding Classification of Gene Expression Data With Feature Selection: A Comparative Study", *International Journal of Computational Intelligence Research*, ISSN 0973-1873 Vol.1, No.1 (2005), pp. 68–76
- [5] Shital Shah, Andrew Kusiak, "Cancer gene search with data-mining and genetic algorithms", *Computers in Biology and Medicine* 37 (2007) 251 – 261.
- [6] Jiexun Li, Hua Su, Hsinchun Chen, *Fellow, IEEE*, and Bernard W. Futscher, "Optimal Search-Based Gene Subset Selection for Gene Array Cancer Classification", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 11, No. 4, July 2007.
- [7] Gregory Piatetsky-Shapiro, Pablo Tamayo, "Microarray Data Mining: Facing the Challenges", *SIGKDD Explorations*, Volume 5, Issue 2 - Page 1.
- [8] Mohd Saberi Mohamad, Sigeru Omatu, Michifumi Yoshioka, Safaai Deris, "An Approach Using Hybrid Methods to Select Informative Genes from Microarray Data for Cancer Classification", *Second Asia International Conference on Modeling & Simulation*, 978-0-7695-3136-6/08, IEEE.

# Effort Estimation Using Use Case Point for Software Development

A. Prof. Svapnil Vakharia, B. Mr. Mukesh Parmar

**Abstract-**This paper describes study of an effort estimation method based on use cases, the Use Case Points method. The original method was adapted to incremental development and evaluated on a large industrial system with modification of software from the previous release. We modified the following elements of the original method: a) complexity assessment of actors and use cases, and b) the handling of non-functional requirements and team factors that may affect effort. The study identified factors affecting effort on projects.

**Keywords:** Estimation, use cases, incremental development.

## 1. INTRODUCTION

Effort estimation is a challenge in every software project. The estimates will impact costs and expectations on schedule, functionality and quality. While expert estimates are widely used, they are difficult to analyze and the estimation quality depends on the experience of experts from similar projects. Alternatively, more formal estimation models can be used. Traditionally, software size estimated in the number of Source Lines of Code (SLOC), Function Points (FP) and Object Points (OP) are used as input to these models, e.g. COCOMO and COCOMO II.

Because of difficulties in estimating SLOC, FP or OP, and because modern systems are often developed using the Unified Modeling Language (UML), UML-based software sizing approaches are proposed. Examples are effort estimation methods based on use cases and software size estimation in terms of FP from various UML diagrams.

The *Use Case Points* (UCP) estimation method introduced in 1993 by Karner estimates effort in person-hours based on use cases that mainly specify functional requirements of a system

Use cases are assumed to be developed from scratch, be sufficiently detailed and typically have less than 10-12 transactions. The method has earlier been used in several industrial software development projects and in student projects. There have been promising results and the method was more accurate than expert estimates in industrial trials

A. Prof. Svapnil Vakharia is with the Department of Information Technology, Gandhinagar Institute of Technology, [svapnil.vakharia@git.org.in](mailto:svapnil.vakharia@git.org.in)

B. Mr. Mukesh Parmar is with the Department of Computer Engineering, Gandhinagar Institute of Technology, [mukesh.parmar@git.org.in](mailto:mukesh.parmar@git.org.in)

Recently, incremental or evolutionary development approaches have become dominant: requirements are covered (or discovered) in successive releases, and changing requirements (and software) is accepted as a core factor in software development.. Each release of a software system may in turn be developed in iterations of fixed or variable duration, and may be maintained for a while, before being replaced with a new release. Project management in these projects needs an estimation method that can estimate the effort for each release based on changes in requirements. Furthermore, software is built on a previous release that should be modified or extended.

## II. THE UNDERLYING ESTIMATION METHODS

### A. The Use Case Points Estimation Method

A use case model defines the functional scope of the system to be developed. Attributes of a use case model may therefore serve as measures of the size and complexity of the functionality of a system.

The Use Case Points (UCP) estimation method is an extension of the *Function Points Analysis* and *MK II Function Points Analysis*. Table 3 gives a brief introduction of the six-step UCP method. In step 4, there are 13 *technical factors* Table 1 (related to how difficult it is to build the system, e.g. distributed system, reusable code and changeability) and eight *environmental factors* Table 2 (related to the efficiency of the project e.g. object-oriented experience and stable requirements). The weights and the formula for technical factors is borrowed from the Function Points method proposed by Albrecht. The formula for environmental factors is based on some estimation results. In step 6, the adjusted Use Case Points (UCP) is multiplied by person-hours needed to implement each use case point .

Table 1. Technical Complexity Factors

Technical Factor	Description	Weight
T1	Distributed system	2
T2	Performance	1
T3	End User Efficiency	1
T4	Complex internal Processing	1
T5	Reusability	1
T6	Easy to install	0.5
T7	Easy to use	0.5
T8	Portable	2
T9	Easy to change	1
T10	Concurrent	1
T11	Special security features	1
T12	Provides direct access for third parties	1
T13	Special user training facilities are required	1

Table 2 Environmental Complexity Factors

Technical Factor	Description	Weight
E1	Familiarity with UML	1.5
E2	Application Experience	0.5
E3	Object Oriented Experience	1
E4	Lead analyst capability	0.5
E5	Motivation	1
E6	Stable Requirements	2
E7	Part-time workers	-1
E8	Difficult Programming language	2

Table 3. The UCP estimation method

Step	Rule	Output
1	Classify actors: a) Simple, WF (Weight Factor) = 1 b) Average, WF = 2 c) Complex, WF = 3	Unadjusted Actor Weights  (UAW) = $\Sigma(\#Actors * WF)$
2	Classify use cases: a) Simple- 3 or fewer transactions, WF = 5 b) Average- 4 to 7 transactions, WF = 10 c) Complex- more than 7 transactions, WF= 15	Unadjusted Use Case Weights  (UUCW) = $\Sigma(\#Use Cases * WF)$
3	Calculate the Unadjusted Use Case Point (UUCP).	$UUCP = UAW + UUCW$
4	Assign values to the technical and environmental factors [0..5], multiply by their weights [-1..2], and calculate the weighted sums (TFactor and EFactor). Calculate TCF and EF as shown.	Technical Complexity Factor (TCF) = $0.6 + (0.01 * TFactor)$ Environmental Factor (EF) = $1.4 + (-0.03 * EFactor)$
5	Calculate the adjusted Use Case Points (UCP).	$UCP = UUCP * TCF * EF$
6	Estimate effort (E) in person-hours.	$E = UCP * PHperUCP$

### III. THE ADAPTED UCP ESTIMATION METHOD

#### A. Overview of the Adapted Method

This section describes how the UCP method has been modified. The new rules are summarized in Table 4, with the same abbreviations as in Table 3 and as described below.

Step 1. Actors. An actor may be a human, another system or a protocol. Since the classification has little impact on the final estimation result, all actors are assumed to be average. Modified actors are also counted as the Modified Unadjusted Actor Weights (MUAW).

Step 2. Counting the UUCW and MUUCW. We started to count the Unadjusted Use Case Weights (UUCW) for Release 1 using the method described in Section 2.1. All use cases in this study would be classified as complex. Use cases should then be classified. Applying step 2 in Table 4 led to most new use cases being classified as simple (66%), and very few as complex. However, the complexity of transactions does not justify such distribution.

Karner proposed not counting so-called included and extended use cases, but the reason is unclear. We have applied the same rules to all the use cases.

Step 3. Counting UUCP. The Unadjusted Use Case Points (UUCP) are calculated; once for all use cases (in Rule 3.1) and once for modifications (in Rule 3.2).

Step 4. TCF and EF. Assigning values to technical and environmental factors are usually done by experts or project leaders, based on their judgment. The impact of the Technical Complexity Factor (TCF) is small and it does not cover all the non-functional requirements either. We have therefore handled this otherwise, as described in step 6 below. The Environmental Factor (EF) is not relevant in this project as there are few changes to this factor from one release to another. However, this factor does have a large impact on the estimate and we have accounted for omitting it by using a high PHperUCP. Dropping these factors is also suggested in other cost models.

Step 5. The adjusted UCP and MUCP will be equal to the unadjusted ones since TCF and EF are set to 1.

Step 6. We assume that there are two mechanisms that consume effort in our model:  $E_{primary}$  estimates effort for realizing new and modified use cases, while  $E_{secondary}$  estimates effort for *secondary changes* as described below. The total estimated effort is the sum of these.

B. Effort Estimation for Secondary Changes of Software

In addition to changes related to functionality estimated in  $E_{primary}$ , there are several other reasons for why software is modified:

1. Perfective functional changes not specified in use cases: Functionality is also enhanced and improved in each release by initiating so-called change requests after requirement freeze. There may also be ripple effects of changes in use cases.
2. Perfective non-functional changes: Quality attributes are improved between releases (performance, security, reliability etc.), but these changes are not reflected in use cases. Improving quality attributes is usually achieved by modifying software that is already implemented.
3. Corrective changes: Some effort is also spent on modifying software to correct detected defects.
4. Preventive changes to improve design and reduce software decay also consume effort. Also note that preventive changes to improve file structure or to reduce dependencies between software modules may later impact quality attributes such as maintainability.

Table 4 The adapted UCP estimation method

Step	Rule	Output
1	1.1. Classify all actors as average, $WF = 2$ .	$UAW = \#Actors * 2$
	1.2. Count the number of new/modified actors.	Modified UAW (MUAW) = $\#New\ or\ modified\ actors * 2$
2	2.1. Since each transaction in the main flow contains one or several transactions, count each transaction as a single use case. 2.2. Count each alternative flow as a single use case. 2.3. Exceptional flows, parameters, and events are given weight 2. Maximum weighted sum is limited to 15 (a complex use case). 2.4. Included and extended use cases are handled as base use cases. 2.5. Classify use cases as: a) Simple- 2 or fewer transactions, $WF = 5$ b) Average- 3 to 4 transactions, $WF = 10$ c) Complex- more than 4 transactions, $WF = 15$	Unadjusted Use Case Weights (UUCW) = $\sum (\#Use\ Cases * WF) + \sum (\#Use\ Case\ Points\ for\ exceptional\ flows\ and\ parameters\ from\ 2.3)$
	2.6. Count points for modifications in use cases according to rules 2.1-2.5 to calculate the Modified Unadjusted Use Case Weights (MUUCW).	$MUUCW = \sum (\#New/modified\ Use\ Cases * WF) + \sum (\text{Points for new/modified exceptional flows and parameters})$
3	3.1. Calculate UUCP for all software.	$UUCP = UAW + UUCW$
	3.2. Calculate Modified UUCP (MUUCP).	$MUUCP = MUAW + MUUCW$

4	Assume average project.	$TCF = EF = 1$
5	5.1. Calculate adjusted Use Case Points (UCP).	$UCP = UUCP$
	5.2. Calculate adjusted Modified UCP (MUCP).	$MUUCP = MUUCP$
6	6.1. Estimate effort for new/modified use cases.	$E_{primary} = MUUCP * PH_{perUCP}$
	6.2. Estimate effort for secondary changes of software.	$E_{secondary} = (UCP - MUUCP) * EMF * PH_{perUCP}$
		$E = E_{primary} + E_{secondary}$

IV. CONCLUSION

From this paper, we classified two techniques of Effort Estimation through Use Case point. Through this evaluate Technical Complexity Factor and Environmental Complexity Factor. It will make very clear and enhanced method for finding out Effort Estimation. One main assumption is that use cases may be used as a measure of the size of a system, and that changes in these may be used as a measure of changes in functionality between releases. Generally, predicting the size of a software system in SLOC or FP in early phases is as difficult as predicting the needed effort. For changes not reflected in use cases, an additional model is used.

The method does not depend on any tools (although there are tools for the original UCP method), paradigms or programming languages, and can promote high quality use cases. The method is cheap, transparent and easy to understand. The method is also suitable when development is being outsourced.

The proposed changes for breaking down use cases and new classification rules were necessary since use cases are written with different levels of details in different projects.

Two mechanisms of effort consumption are identified:

- a) Primary changes reflected in changes in use cases, and
- b) Secondary changes or modification of software from a previous release as a ripple effect of primary changes, unplanned changes and improvements in quality attributes.

Use case diagrams are usually available before other UML diagrams, but have variable level of details. A challenge is to define some standards for these. Furthermore, use cases essentially express functional requirements. The influence of nonfunctional requirements should be included in the technical factors, the number of PHperUCP or as the estimated effort for secondary changes.

There are few empirical studies on estimation of incrementally developed projects. The UCP method showed flexibility in adapting to the context, but there are many project-specific factors in the original method and in our extensions of it. Future studies can help to understand the degree of modification in incremental development of a system (use cases, code and integration costs), and how the method works on other types of systems.

**REFERENCES**

- [1]Boehm, B., Clark, B., Horowitz, E., Westland, C., Madachy, R., Selby, R. Cost models for future software life cycle processes:” *COCOMO 2.0. USC center for software engineering,*” information retrieved on 01/03/2012 from, <http://sunset.usc.edu/publications/TECHRPTS/1995/index.html>
- [2]Schneider, G., Winters, J.P. *Applying Use Cases, a Practical Guide.* Addison-Wesley, 1998.
- [3] Smith, J. The estimation of effort based on use cases.*Rational Software,* White paper, 1999.
- [4] Symons, P.R. *Software Sizing and Estimating MK II FPA(Function Point Analysis).* John Wiley & Sons, 1991.
- [5]Kemerer, C.F. An empirical validation of software cost estimation models. *CACM*, 30, 5 (May 1987), 416- 429

# Comparative Evaluation and Selection of Grid Computing Middleware

A. Nisha Patel

*Abstract—It has become a known fact that the rate at which Grid Computing is rising nowadays in terms of applications and demand both in research institutes and industries which need not be over emphasized. The ubiquitous nature of this new innovation and technology coupled with open source nature of most of the softwares have led to the research in various Grid Middlewares; without which grid resources cannot be shared. These middlewares allow submission of requests to execute a computation (called a Job) to the Grid, such that it can be run anywhere on the network. Therefore, grid Middlewares serve as an intermediary layer that allow a reliable and homogeneous access to resources managed locally with different syntax and access techniques. Within the context of availability of various Middlewares for Grid implementation with different features, this paper focuses on various features that are peculiar to Alchemi by taking into consideration its Architecture, the Operating systems, software demand and limitation that are inherent from its usage.*

## I. INTRODUCTION

THERE has been a lot of work on Grid Computing already carried out and still being researched upon in various institutions and laboratories across the globe. Work done in evaluating various grid framework options and choosing the appropriate one for given project is briefly outlined here. The chosen framework and its architecture are detailed in the later part emphasizing the rightfulness of its choice. Laboratory scale experimental implementation of the framework has been evidenced in the end proving its effectiveness as the proper option.

### A. PVM

The Parallel Virtual Machine (PVM) was the former de facto standard for parallel computer programming. PVM programs are better in communicating between independent programs, but MPI is somewhat faster. PVM also ensures that different languages, such as C and Fortran, can communicate with each other. [1]

A. Author is with the Gandhinagar Institute of Technology (e-mail: [nisha.patel@git.org.in](mailto:nisha.patel@git.org.in))

### B. MPI

The Message Passing Interface (MPI) is the current de facto standard for parallel computer programming. It is based on a distributed shared memory model. MPI is programming language agnostic but most implementations only have directly callable routines for Fortran, C and C++. Other programming libraries can use the MPI indirectly using libraries. Well known software implementation of MPI include MPICH2, Open MPI and OpenMP.

### C. Alternatives for MPI

In 1996, PVM (Parallel Virtual Machine) was the de facto standard for distributed computing [1], while MPI was still new in the field. PVM and MPI both had their own merits, but MPI was the result of a standardization effort and has more implementations [2]. For more alternatives and comparisons, see [1] and [2].

### D. Globus

The Globus Toolkit is used to provide an infrastructure for grid computing. This involves the management of resources such as nodes on the network, users, and security. Globus also provides Web Services [3].

### E. Condor

Condor is a batch system backend for grids. Think job queuing, scheduling, priority management and so on. It wraps around Globus and provides an infrastructure for multi-institutional grids. Condor-G automates the GRAM functions of Globus by, for example, tracking submitted jobs and handling job failures.

### F. gLite

gLite is grid middleware produced by the EGEE (Enabling Grids for E-science), a project sponsored by the European Union. The EGEE project ended on 2010-04-30, but the grid infrastructure (including gLite) is now maintained by the European Grid Infrastructure. gLite makes use of Globus and Condor software.

### G. SAGA

SAGA stands for “Simple API for Grid Applications” and does not replace Globus or other grid middleware. It is meant for application programmers that have no experience with



grid programming and want to develop grid-enabled applications without much difficulty.

#### H. BOINC

A different approach is taken by the Berkeley Open Infrastructure for Network Computing (BOINC). This is also a middleware grid system. It originates from the SETI@home project. A very interesting feature of BOINC is its use of Graphical Processing Units (GPUs), better known as 3D graphics cards. It supports the CUDA programming interface of Nvidia.

#### I. Apple Xgrid

Apple Xserve supports multiple MPI implementations. It uses Mac OS X clients as worker nodes. An Xgrid client is pre-installed on all OS X versions beginning with version 10.4.

#### J. Microsoft HPC

Microsoft has released the Windows HPC<sup>1</sup> Server 2008 R2 with its own MPI implementation called “Microsoft MPI”, which is based on MPICH2 but has a few Windows specific features, such as Windows integrated security. A free SDK is available. The server can make use of worker nodes, but client software must be installed on the workers. Limited interoperability with Linux MPI systems is in place, and also there are tools to make use of grid computing in Excel spreadsheets.

#### K. MPI.NET

Although MPI.NET is just another MPI enabler, using other implementations, it is worth being mentioned here. MPI.NET is developed by the Indiana University to enable MPI in the .Net programming environment, including Mono. When used on Unix / Linux (\*nix) platforms, it compiles under Mono and uses existing nix MPI implementations. When installed on Windows, it uses the Microsoft MPI implementation.

#### L. Alchemi

Alchemi is open-source software for Windows .Net that provides its own methods for parallel computing and its own worker (“executor”) and broker (“manager”) programs. Its performance is somewhat inferior to the Globus Toolkit [4]. It was developed by the Grid Computing and Distributed Systems (GRIDS) Laboratory of the University of Melbourne [5]. Alchemi can interoperate with other grid middleware and other platforms using its optional web service [6].

#### M. SGP(A functional programming approach)

Simple Grid Protocol (SGP) is a set of protocols and software for Linux and BSD. It is programmed in CLISP (a Common Lisp implementation). LISP is a well-known functional programming language. Each function can be executed in parallel. SGP is a small and not widely used project, but the functional approach is interesting.

#### N. Linux distributions

A number of Linux distributions exists that offer a one-in-all installation of the whole software chain (operating system, platform, grid middleware, mpi libraries, compilers, source

editors, browsers, and other tools). Most of them can run from a Live CD or are USB bootable.

#### Grid Appliance

The Grid Appliance is a virtual appliance that can be run using VMware or VirtualBox. It is a fast and easy way to deploy a test and development Condor / MPI grid for education and training. It is self-configuring and works with peer-to-peer VPN networks. You can run multiple instances of the Grid Appliance on the same computer, but it is also possible to connect these instances with other instances on other computers. Grid Appliance is based on Ubuntu.

#### Instant Grid

Instant Grid is similar to the Grid Appliance, but comes as a Live CD. It is based on Knoppix, which itself is Debian based. The first computer booted with the CD acts as a DHCP and PXE server, so other nodes (workers) can easily be added to the grid. The worker nodes are booted over the network and are automatically configured. The grid middleware is formed by the Globus Toolkit 4. It comes with a lot of additional software, such as the Ganglia grid monitoring system and Gobby for collaborative editing. Instant Grid was developed by some German institutions, and the project is now closed. Although development has halted, some participants in Germany are still using it.

#### PelicanHPC

PelicanHPC (former name: Parallel Knoppix) is very similar to Instant Grid. It has many of the same features, such as a single boot CD with a DHCP and PXE server. It is now based on the Debian Live base [7]. BirgHPC (Bioinformatic Research Group High Performance Computing) is a derivative that includes additional software such as MPICH2 and a bioinformatics tool GROMACS.

#### Scientific Linux

Scientific Linux is a compilation of Red Hat Enterprise Linux with many scientific applications. It is maintained by Fermilab and CERN. gLite uses Scientific Linux as the primary method for quick deployments and testing.

## II. SELECTION OF THE GRID FRAMEWORK

The selection of grid framework, amongst other, involved two important and determining considerations (1) Choice of the OS: Proprietary or Open Source, (2) Choice of Middleware: Proprietary or Open Source.

In present case, the OS choice was easy to make. Though proprietary, Microsoft Windows XP was chosen because most of the machines involved in this work already had it. In this way making choice of a proprietary OS was not going to attract any additional financial aid and it was even more convenient to find support on the same from available sources.

Having chosen OS, the middleware working on the selected OS platform were to be evaluated. Further they need to be open source for minimizing any additional need for financial

aid. This ended the process at selection of Alchemi. Going with Windows software, the Alchemi framework is based on .NET and is easy to install, administer, and use.

### III. ALCHEMI API

Alchemi offers a powerful grid thread programming model. One can develop, execute and monitor grid applications using the .NET API and tools.

Many of the Software to enable grid computing have been primarily written for Unix-class operating systems, thus severely limiting the ability to effectively utilize the computing resources of the vast majority of desktop computers i.e. those running variants of the Microsoft Windows operating system. Addressing Windows-based grid computing is particularly important from the software industry's viewpoint where interest in grids is emerging rapidly.

Microsoft's .NET Framework has become near-ubiquitous for implementing commercial distributed systems for Windows-based platforms, positioning it as the ideal platform for grid computing in this context. Alchemi, a .NET-based grid computing framework provides the runtime machinery and programming environment required to construct desktop grids and develop grid applications. It allows flexible application composition by supporting an object-oriented grid application programming model in addition to a grid job model. Cross-platform support is provided via a web services interface and a flexible execution model supports dedicated and non-dedicated (voluntary) execution by grid nodes.

There are four types of distributed components (nodes) involved in the construction of Alchemi grids and execution of grid applications: Manager, Executor, User & Cross-Platform Manager.

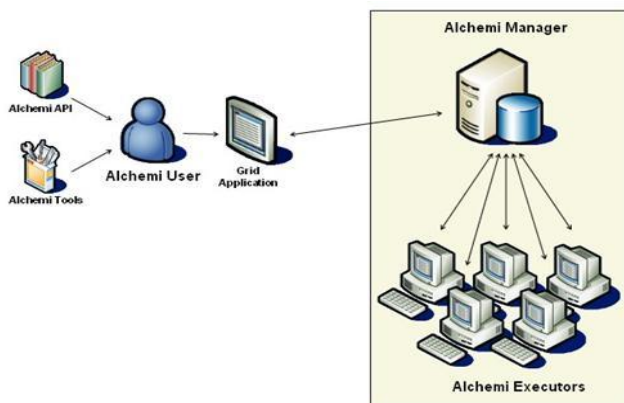


Figure 1. An Alchemi Grid

A grid is created by installing Executors on each machine that is to be part of the grid and linking them to a central Manager component.

An Executor can either be dedicated or non-dedicated. A non-dedicated Executor works along NAT servers and firewalls because there is a one-way interaction and communication

between the Manager and the Executor. Dedicated Executors are very useful and more applicable in an intranet environment while internet environment is mostly suitable for non-dedicated Executors.

Grid application using Alchemi can be monitored and executed using the .NET API and with other tools which are part of Alchemi SDK. Alchemi provides grid thread programming model which simplifies the construction of grid applications and a grid job model for non-.Net applications.

The Cross-Platform Manager, an optional component (not shown in figure 1) is the web service which offers interoperability with custom non-.NET grid middleware. This is purely a web based service interface that reveals a section of the functionality of the Manager so as to enable Alchemi manage the job execution on the grid system. Jobs submitted to the Cross-Platform Manager are converted into a form that is understandable and acceptable to the Manager.

User is any application to be executed on the grid Alchemi, usually done on the user node. The API conceptualizes the grid implementation from the user and performs different services such as submitting application and its constituent threads for job execution. Consequently, it notifies the user of completed job and also informs him of failed threads along with reasons for the said errors.

The Manager manages any grid application(s) and their constituent threads. Executors monitor their status after they have already registered with the manager. As threads are being received from the users they are also placed in a pool and scheduled for execution in various Executors. A priority is placed on the thread when any job is created and submitted. Threads are scheduled on a Priority and First Come First Served (FCFS) basis, in that order. The completed threads are usually returned back to the Manager by the Executors and these jobs will be received by their respective users [4].

### IV. ALCHEMI ARCHITECTURE

Alchemi follows the master-worker parallel programming paradigm in which a central component dispatches independent units of parallel execution to workers and manages them. This smallest unit of parallel execution is a grid thread, which is conceptually and programmatically similar to a thread object that wraps a "normal" multitasking operating system thread. A grid application is defined simply as an application that is to be executed on a grid and that consists of a number of grid threads. Grid applications and grid threads are exposed to the grid application developer via the object oriented Alchemi .NET API.

A scenario for the creation of a layered architecture-based grid computing environment using Alchemi Toolkit is shown in Figure2.

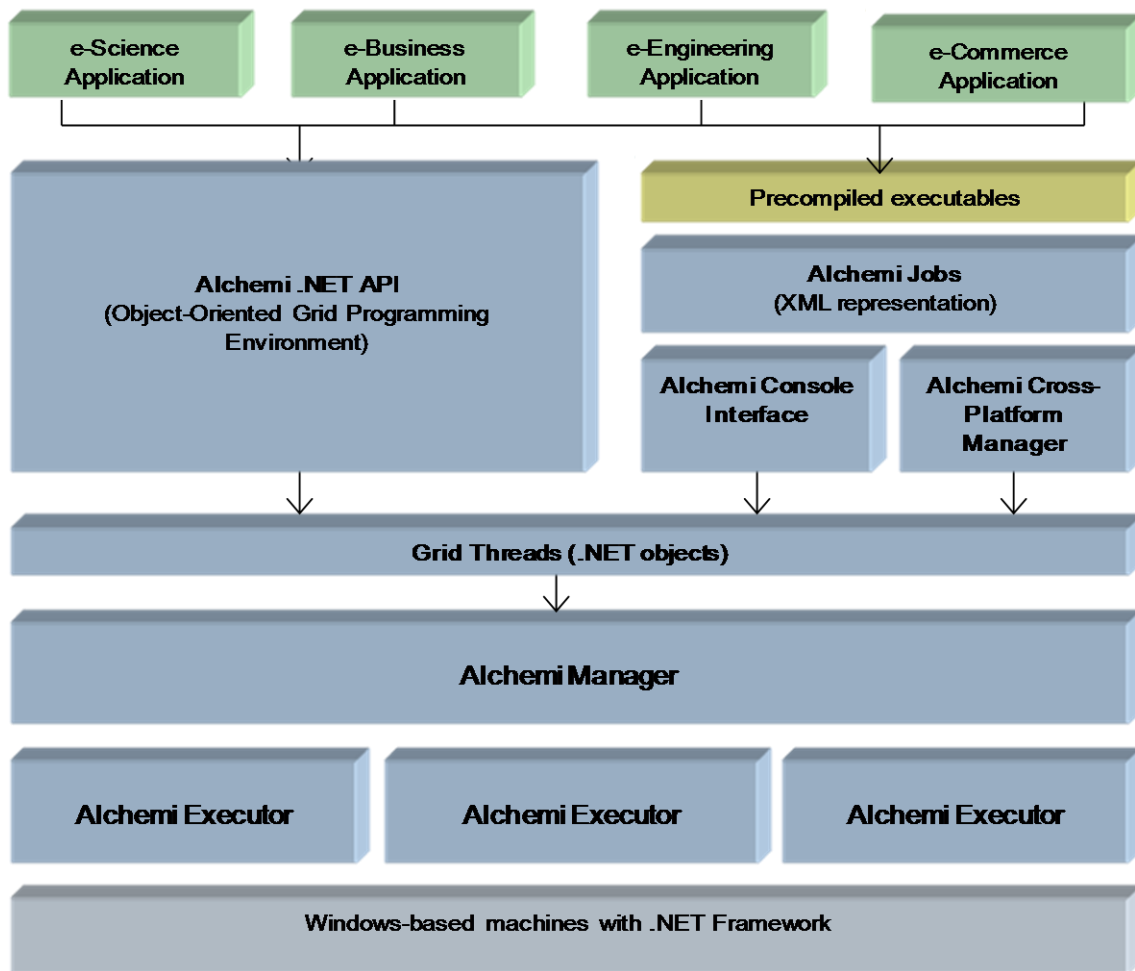


Figure 2. Alchemi layered architecture

TABLE 1: TIME TAKEN BY EXECUTOR TO COMPLETE TASK

EXECUTORS QUANTITY	1	2	3	4	5
PI Digits(100)	78	74	72	68	62
PI Digits(200)	97	89	83	76	70
PI Digits(300)	131	106	92	84	74
PI Digits(400)	136	114	100	90	76
PI Digits(500)	143	123	112	95	82

V. TEST APPLICATION & METHODOLOGY

The application utilizes the Alchemi grid thread model .The test application is the computation of the value of Pi to n decimal digits. The algorithm used allows the computation of the p'th digit without knowing the previous digits. The test was performed for a range of workloads (calculating 100, 200, and 300 ,400 & 500 digits of Pi),each with one to Five Executors enabled. The workload was sliced into a number of threads, each to calculate 20 digits of Pi, with the number of threads varying proportionally with the total number of digits to be calculated. Execution time was measured as the elapsed clock time for the test program to complete on the Owner node. If the workload is low, there is a little difference between the total execution time with different quantity of executors. As the workload is increased, the total overhead is

in a relatively lower proportion to the actual total completion time.

## REFERENCES

- [1] G. A. Geist, J. A. Kohl, and P. M. Papadopoulos. Pvm and mpi: a comparison of features, 1996.
- [2] Jack Dongarra and Alexey Lastovetsky. An overview of heterogeneous high performance and grid computing.
- [3] Ian Foster. A Globus Primer - Or, Everything You Wanted to Know about Globus, but Were
- [4] Afraid To Ask - Describing Globus Toolkit Version 4 (draft), 2005.
- [5] Ahsan Arefin, Shiblee Sadik, Serena Coetzee, and Judith Bishop. Alchemi vs globus: A performance comparison. Electrical and Computer Engineering, (19-21 Dec.):602–605, 2006. 4th International Conference on Electrical and Computer Engineering.
- [6] Alchemi: A .NET-based enterprise grid computing system, 2006.
- [7] Krishna Nadiminti and Rajkumar Buyya. Enterprise grid computing: State of the art, 2005.
- [8] Michael Creel. PelicanHPC.  
<http://pareto.uab.es/mcreel/ParallelKnoppix/>.
- [9] <http://sharptoolbox.com/categories/grid-computing.>: accessed on 24th June,2010

# SURVEY OF DATA PREPROCESSING FOR WEB LOG

A. Brinda Parekh, B. Pooja Mehta

**Abstract**—With the Internet usage gaining fame and the stable growth of users, the World Wide Web has become a huge repository of data. Web mining is the application of the data mining which is useful to extract the knowledge/information from Internet. Web usage mining, Web structure mining and Web content mining are the types of Web mining. Web usage mining is used to mining the data from the web server log files, Client side cookies, software agent etc. Before analyzing such data using web mining techniques, the web log has to be pre processed, integrated and transformed. Today's real world databases are highly susceptible to noisy, missing and inconsistent data due to their typically huge size data and their origin from multiple, heterogeneous sources. Hence, pre-processing of data is necessary to help improve the quality of data and consequently the mining results. There are many data pre-processing techniques. In this paper, we would like to discuss different approaches for data pre-processing specifically for web log files

**Index Terms**—Data Pre-processing, User Identification, Session Identification, Web logs, Web usage mining.

## I. INTRODUCTION

The World Wide Web has become one of the most important medium to store, share information. The discovery and analysis of useful information from the Web documents is referred to as Web mining [2]. But the data stored in the Web is heterogeneous. Web data are classified into four categories: content (simple text, images), structure (html, xml tags), usage (visitor's ip address, time, date of access, url etc), and user profiles (information about user's preferences, interest etc). Due to the properties of the huge, diverse, dynamic and unstructured nature of Web data, raise great challenge to web user to extract useful information. Thus preparation of a proper data set is pre-requisite and an important task for mining the web data. Hence the web usage mining process is divided into three phases as shown in Fig. 1: data pre-processing, pattern discovery, pattern analysis.

A. Author is with the Gandhinagar Institute of Technology, Moti Bhojan, and Gujarat, India ( e-mail: brinda.parekh@git.org.in).

B. Author, is with Sabar Institute of Technology for Girls (e-mail: poojamehta810@gmail.com ).

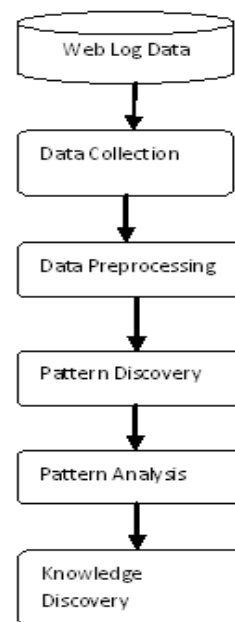


Fig.1. Web Usage mining process

Data Pre-processing basically consists of data cleaning, data integration, transformation and data reduction. If there is much irrelevant and redundant information present or noisy and unreliable data, then by using data pre-processing we can clean and filter the data. This technique includes cleaning, normalization, transformation, feature extraction and selection etc. Pattern discovery deals with extracting knowledge from pre-processed data. Some of the techniques used in Pattern discovery are Association rules, Classification, Clustering etc. Pattern Analysis filters out uninteresting rules or patterns from the set found in the pattern discovery phase [1].

Following are the task in the data preprocessing techniques.

- Data cleaning
- Data transformation
- Attribute / feature construction
- Data reduction and discretisation
- Data parsing and standardisation
- Data integration and linkage

The data pre-processing is the most time consumption phase of the web page analysis [10]. In data pre-processing we measure the following attribute of the data.

- Accuracy

- Completeness
- Consistency
- Timeliness
- Believability
- Interpretability
- Accessibility

The Web data is stored in Web servers, client machines, proxy servers or organizational databases. The primary data sources used in Web usage mining are the server log files which include Web server access logs, referrer logs and agent logs. Additional data sources that are also essential include the site files and meta-data, operational databases and domain knowledge [2]. Data pre-processing is predominantly significant phase in Web usage mining due to the properties of Web data. This phase is often the most time-consuming and computationally intensive step in Web usage mining. This process is critical to the success of Pattern discovery and Pattern Analysis.

In this paper, we would be focusing on different approaches of data pre-processing. The rest of this paper is organized as follows: Section II overview of web log file, Section III gives basic steps involved in pre-processing phase, Section IV reviews related work, Section V gives the future work and lastly Section VI gives the concluding note.

## II. WEB LOG

A web server log file is a simple plain text file which records information each time a user requests a resource from a web site. Web log files provide web administrators with many useful kind of information like [6]:

- Which pages of your web site were requested
  - What are the errors that people encounter
  - What is the status returned by the server upon user request
  - How many bytes sent from the server to the user
- Generally there are four types of server logs:[6]
- *Access log file*- contains data of all incoming requests and lets you track and get information about clients of the server.
  - *Error log file*- lists internal server errors, enables server administrators to correct site content or to detect anomalous activities.
  - *Agent log file*- provides information about user's browsers, operating system and browser version.
  - *Referrer log file*- provides information about the link that redirects visitors to site.

An example of one entry in a log file: The following is a fragment from the IIS server logs: [1]

```
date time c-ip cs-username s-sitename s-computername sip
s-port cs-method cs-uri-stem cs-uri-query sc-status timetaken
cs-version cs-host cs(User-Agent) cs(Referer)
```

```
2010-11-19 19:24:24 W3SVC1 192.168.1.57
GET/sharedoutandabout/InAndOut/ Categories.aspx
insid=2&langid=1 80 - 93.186.23.240 Mozilla/4.0+
(compatible;+MSIE+4.01;+Windows+NT) 200 3223
```

The interpretation of the fields: [6]

- 1) 2010-11-19: Date at which the entry is recorded.
- 2) 19:24:24: Time at which the entry is recorded
- 3) W3SVC1: Name of the server
- 4) 192.168.1.57: IP address of the server
- 5) Get: Method of HTTP request
- 6) sharedoutandabout/InAndOut/Categories.aspx: The resource requested
- 7) insid=2&langid=1: Parameters associated with the resource requested
- 8) 80: Port number
- 9) -: This is the user name if the site require user authentication. If not the hyphen is placed
- 10) 93.186.23.240: Client IP address
- 11) Mozilla/4.0+ (compatible;+MSIE+4.01;+Windows+NT): Browser name and version and the operating system.
- 12) 200: It is the status code returned to the user which in this case means that the request is successfully executed.
- 13) 3223: The bytes sent from the server to the client in response to the user request.

There are different log file format that differ in the number of parameters and fields recorded in each file and in the format of these parameters. The most commonly used format are: NCSA common, NCSA combined, W3C extended format and IIS format.

A log file can be located in three different places [9]: 1) Web server 2) Web proxy server 3) client browsers.

These suffer from two major limitations. Server side logs contain sensitive and personal information therefore the server owner usually keeps them closed. In proxy – server construction is a difficult task .for that advance network programming such as TCP/IP is required for this construction. In client – side logs ,which makes it hard to achieve compatibility with a range of operating system and web browsers.

## III. BASIC STEPS INVOLVED IN PRE-PROCESSING

Following are steps involved in pre-processing. The Steps are: [1][4]

**1) Data Cleansing:** Irrelevant records are eliminated during data cleansing. Since target of web usage mining is to get traversal pattern, following two kinds of records are unnecessary and should be removed.

**a. The records of graphics, video and forma information.**

The records having filenames suffixes of GIF, JPEG, CSS and so on, which can be found in cs\_uri\_stem field of record.

**b. The records with failed HTTP status codes.** By examining the status field of every record in the web log, The record with status code over 299 and under 200 are removed.

**2) User and Session Identification:** The task of user and session identification is to find out the different user sessions from the original web access log. A referrer-based method is used for identifying sessions. The different IP addresses distinguish different users.

*a. If the IP addresses are same.* The different browsers and operation systems indicate different users which can be found by client IP address and user agent who gives information of user's browsers and operating system.

*b. If all of the IP address, browsers and operating systems are same.* The referrer information should be taken into account. The Refer URI (cs\_referer) is checked, a new user session is identified if the URL in the Refer URI - field hasn't been accessed previously, or there is a large interval (usually more than 30 minutes) between the accessing time of this record.

**3) Path Completion:** Path Completion should be used acquiring the complete user access path. The incomplete access path of every user session is recognized based on user session identification. If in a start of user session, Referrer as well URI has data value, delete value of Referrer by adding '-'. Web log pre-processing helps in removal of unwanted click-streams from the log file and also reduces the size of original file by 40-50%.

#### IV. RELATED WORK

In this section we had discussed related work in data pre-processing.

In [1], the authors have discussed two different approaches for pre-processing of the web log file: pre-processing using XML and pre-processing using Text file. In the first approach as suggested by the authors pre-processing is done using XML which provides a structure to the records in the web log file. Logs recorded in the web log which is a text file are converted into DOM tree structure using XML parsers. Followed by this step, user identification and session identification are done. In the second approach web log file are cleaned using text file. Before applying cleansing process, attributes in the text file needs to be separated using delimiter as space. These spaces help in identifying exact position of attributes/fields.

In [2], the author have suggested an approach for data pre-processing of web log file by removing irrelevant items, identify users and sessions along with their browsing information. The output of this phase will result into user session file.

In [3], the suggested pre-processing methodology reduces significantly the size of the initial log files by eliminating unnecessary requests and increases their quality through better structuring.

In [4], the author has given some guidelines and rules for each phase in data pre-processing: data cleaning is to remove the irrelevant and redundant log entries for the mining process. There are three kinds of irrelevant or redundant data needed to clean: accessorial resources embedded in HTML file, robots' requests and error requests. User identification is greatly complicated by the existence of local caches, corporate firewalls, and proxy servers. Rules for this are:

- Each IP address represents one user;
- For more logs, if the IP address is the same, but the agent log shows a change in browser software or operating system, an IP address represents a different user ;
- Using the access log in conjunction with the referrer logs and site topology to construct browsing paths for each user. If a page is requested that is not directly reachable by a hyperlink from any of the pages visited by the user, there is another user with the same IP address.

Goal of session identification is to divide the page accesses of each user into individual sessions. The rules for this phase are:

- If there is a new user, there is a new session;
- In one user session, if the refer page is null, there is a new session;
- If the time between page requests exceeds a certain limit (30 or 25.5minutes), it is assumed that the user is starting a new session.

Path completion, task is to fill in these missing page references. Methods similar to those used for user identification can be used for path completion. If a page request is made that is not directly linked to the last page a user requested, the referrer log can be checked to see what page the request came from. If the page is in the user's recent request history, the assumption is that the user backtracked with the "back" button available on most browsers, calling up cached versions of the pages until a new page was requested. If the referrer log is not clear, the site topology can be used to the same effect. If more than one page in the user's history contains a link to the requested page, it is assumed that the page closest to the previously requested page is the source of the new request.

In [5], the presented approach concentrated on criticality of web log pre-processing. The proposed approach improves the efficiency of pre processing of web log, it separates human user and web search engine accesses automatically, in less time, it reduces the error rate of learning algorithm The work ensures the goodness of split by using popular measures like Entropy and Gini index.

In [8], the author suggested preprocessing methodology which divides the preprocessing method in to three categories. In that the first is web site preprocessing then second in server log preprocessing and the last one is Access information collection. Initially the first method preprocess the data and gives the cleaned web log file then second it will generate the session id and also provide the time spent on each page by user so by this we can get the high quality of the data structure.

In [10], the author try to find out the experiment to what measure it is necessary to execute the data preparation for the web log mining and determine inevitable steps for gaining valid data from the log file. The author tried to find out that to assess the impact of reconstruction of the activities of web visitor on the quantity and quality of the extracted rules which represent the web users' behavior patterns.

## V. FUTURE WORK

As suggested by many authors there are a number of issues in pre-processing of log data.

Volume of requests in web log in a single log file is the first challenge. Analyzing web user access log files helps to understand the user behaviours in web structure to improve the design of web components and web applications. Efforts have to be done to find accurate sessions which are likely to be the most productive in the formation of much effective web usage mining and personalization systems. By following data preparation steps, it is very easier to generate rules which identify directories for website improvement. More research can be done in pre-processing stages to clean raw log files, and to identify users and to construct accurate sessions.

## VI. CONCLUSION

In this paper, we have discussed about the web log files. In the later section discussion about the basic steps involved in data pre-processing have done and lastly different approaches for data pre-processing have been discussed.

## REFERENCES

- [1] Ms. Dipa Dixit, Ms. M Kiruthika, "Preprocessing Of Web Logs", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 07, 2010, p.2447-2452.
- [2] C.P. Sumathi, R. Padmaja Valli, T. Santhanam, "An Overview Of Preprocessing Of Web Log Files For Web Usage Mining", Journal of Theoretical and Applied Information Technology Vol. 34 No.1, 2011, p. 88.
- [3] Ramya C, Kavitha G, "An Efficient Preprocessing Methodology for Discovering Patterns and Clustering of Web Users using a Dynamic ART1 Neural Network", Fifth International Conference on Information Processing, August-2011
- [4] Li Chaofeng, "Research and Development of Data Preprocessing in Web Usage Mining"
- [5] V.V.R. Maheswara Rao, Dr. V. Valli Kumari, "An Enhanced Pre-Processing Research Framework For Web Log Data Using A Learning Algorithm", Netcom, 2011, p.01-15.
- [6] Shaimaa Ezzat Salama, Mohamed I. Marie, Laila M. El-Fangary & Yehia K. Helmy, "Web Server Logs Preprocessing for Web Intrusion Detection", Published by Canadian Center of Science and Education, Vol. 4, No. 4; July 2011, p.123.
- [7] V.Chitraa, Dr. Antony Selvdoss Davamani, "A Survey on Preprocessing Methods for Web Usage Data", International Journal of Computer Science and Information Security, Vol. 7, No. 3, 2010, p.78.
- [8] Yongjian Fu, Ming-Yi Shih, Mario Creado, "Reorganizing Web sites based on User Access pattern", International Journal of Intelligent Systems in Accounting, Finance and Management 11, 2002, pp 39-53.
- [9] K.R Suneetha and Dr.R. Krishnamurthy, "Identifying User behaviour by analyzing Web Server Access log files", International Journal of Computer Science and Network security, Vol.9 No.4, April 2009.
- [10] Michal Munk, Jozef Kapusta, Peter Svec, "Data Preprocessing Evaluation for Web Log mining: Reconstruction Of Activities of Web Visitor", International Conference on computer Science, ICCS 2010.



# Cheater Detection and Cheating Identification based on Shamir Scheme

A. Ms. Rupali Kolambe, B. Prof. Megha Kamble

**Abstract**-In cryptography, a secret sharing scheme is a method for distributing a secret amongst a group of participants, each of which is allocated a share of the secret. The secret can only be reconstructed when the shares are combined together; individual shares are of no use on their own.

The study of secret sharing schemes was independently initiated by Shamir and Blakely in 1979. Since then several other secret sharing schemes were introduced.

When shareholders present their shares in the secret reconstruction phase, dishonest shareholder(s) (i.e. cheater(s)) can always exclusively derive the secret by presenting faked share(s) and thus the other honest shareholders get nothing but a faked secret. Cheater detection and identification are very important to achieve fair reconstruction of a secret.

Our proposed scheme uses the shares generated by the dealer to reconstruct the secret and, at the same time, to detect and identify cheaters

Our proposed scheme is an extension of Shamir's secret sharing scheme.

**Index Terms**- Attacks, Consistency, Detection, Identification, Majority voting, Secret sharing scheme

## I. INTRODUCTION

Shamir's  $(t, n)$ -SS scheme is very simple and efficient to share a secret among  $n$  shareholders.

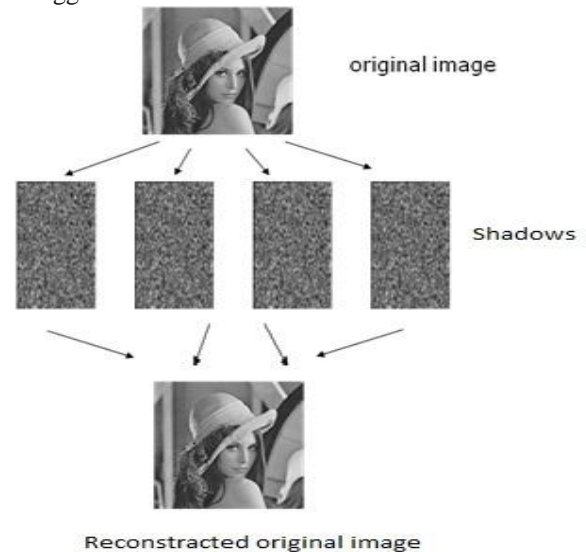
However, when the shareholders present their shares in the secret reconstruction phase, dishonest shareholder(s) (i.e. cheater(s)) can always exclusively derive the secret by presenting faked share(s) and thus the other honest shareholders get nothing but a faked secret.

It is easy to see that the Shamir's original scheme does not prevent any malicious behavior of dishonest shareholders during secret reconstruction.

Cheater detection and identification are very important to achieve fair reconstruction of a secret.

In this paper, we use a different approach to prevent cheaters. We consider the situation that there are more than  $t$  shareholders participated in the secret reconstruction. Since there are more than  $t$  shares (i.e. it only requires  $t$  shares) for reconstructing the secret, the redundant shares can be used for cheater detection and identification. Our proposed scheme uses the shares generated by the dealer to reconstruct the secret and, at the same time, to detect and identify cheaters. Simmons [11] has suggested to use the same method to detect cheaters.

Our proposed scheme uses the shares generated by the dealer to reconstruct the secret and, at the same time, to detect and identify cheaters. Simmons [11] has suggested to use the same method to detect cheaters.



The rest of this paper is organized as follows. In the next section, we provide some preliminaries.

Detection and identification of cheaters we describe attacks of cheaters. We analyze our scheme under three attacks and calculate bounds of detectability and identifiability of our proposed scheme.

A  $(k, n)$  threshold scheme has the following characteristics :

- (1) The secret is divided into  $n$  shadows.
- (2) Any  $k$  or more shadows can be used to reconstruct the secret.
- (3) Any  $k - 1$  or less shadows reveal no knowledge about the secret.

Shamir [1] introduced an elegant and efficient  $(k, n)$

## II. PRELIMINARIES

In this section, we introduce some basic preliminaries.

A. Sagar Institute and Research,  
rupalibe@gmail.com

Bhopal e-mail:

B. Assitant Professor Author, Sagar Institute and Research, Bhopal  
(e-mail: meghakamble@gmail.com).

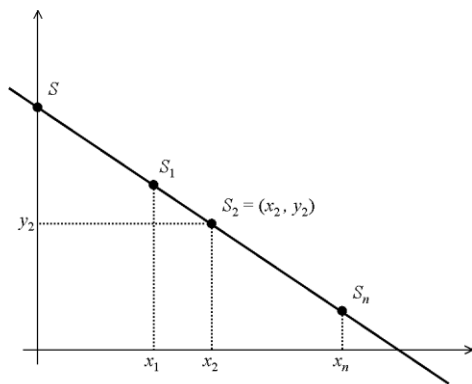
A. Shamir's Secret sharing scheme

Shamir's secret sharing scheme [Sha79] is a threshold scheme based on polynomial interpolation. To allow any  $m$  out of  $n$  people to construct a given secret, an  $(m-1)$ -degree polynomial

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{m-1} x^{m-1}$$

over the finite field  $GF(q)$  is constructed such that the coefficient  $a_0$  is the secret and all other coefficients are random elements in the field; the field is known to all participants. Each of the  $n$  shares is a pair  $(x_i, y_i)$  of numbers satisfying

$f(x_i) = y_i$  and  $x_i \neq 0$ . Given any  $m$  shares, the polynomial is uniquely determined and hence the secret  $a_0$  can be computed. However, given  $m-1$  or fewer shares, the secret can be any element in the field. Therefore, Shamir's scheme is a perfect secret sharing scheme



Shamir's secret sharing scheme

A special case where  $m = 2$  (that is, two shares are required for retrieval of the secret) is given in Figure. The polynomial is a line and the secret is the point where the line intersects with the  $y$ -axis. Namely, this point is the point  $(0, f(0)) = (0, a_0)$ . Each share is a point on the line. Any two points determine the line and hence the secret. With just a single point, the line can be any line that passes through the point, and hence the secret can be any point on the  $y$ -axis

$C$  denotes number of fake shares and  $j(n \geq j \geq t)$  denotes number of participants.  $J = \{i_1, \dots, i_j\}$  algorithms:

1. *Share generation algorithm* the dealer  $D$  first picks a polynomial  $f(x)$  of degree  $t-1$  randomly:

$f(x) = a_0 + a_1x + \dots + a_{t-1}x^{t-1}$ , in which the secret  $s = a_0$  and all coefficients  $a_0, a_1, \dots, a_{t-1}$  are in a finite field  $F$ , and  $D$  computes:

$$s_1 = f(1), s_2 = f(2), \dots, s_n = f(n).$$

Then, the algorithm outputs a list of  $n$  shares  $(s_1, s_2, \dots, s_n)$  and distributes each share  $s_i$  to corresponding shareholder  $P_i$  secretly.

2. *Secret reconstruction algorithm* this algorithm takes any  $t$  shares  $(s_{i_1}, \dots, s_{i_t})$  where  $\{i_1, \dots, i_t\} \subset \{1, 2, \dots, n\}$  as inputs, and outputs the secret  $s$ .

Above scheme satisfies the basic requirements of secret sharing scheme as follows: (1) With knowledge of any  $t$  or more than  $t$  shares, it can reconstruct the secret  $s$  easily; (2) With knowledge of fewer than  $t$  shares, it cannot get any information about the secret  $s$ . Shamir's scheme is *information-theoretically secure* since the scheme satisfies these two requirements without making any computational assumption.

B. Secrets majority

If the shares  $s_1, \dots, s_m$  are *inconsistent*, it is easy to see that secrets  $s_i$  for  $i = 1, \dots, u$  reconstructed by combinations of  $t$  out of  $m$  shares are not identical.

Then, we can divide the set  $U = \{s_1, \dots, s_u\}$  containing all reconstructed secrets into several mutually disjoint subsets  $U_i$ , for  $i = 1, \dots, v$ . Each subset contains same secret.

These subsets satisfy following conditions.

$$U = U_1 \cup \dots \cup U_v, \text{ where } U_i = \{s^{i1}, \dots, s^{iwi}\} \text{ and } s^{wi} = s^{i1} = \dots = s^{iwi};$$

$$U_k \cap U_l = \emptyset \text{ for } 1 \leq k, l \leq v \text{ and } k \neq l.$$

For all subsets  $U_i$  for  $i = 1, \dots, v$  as defined previously, set  $w_i = |U_i|$  and  $w_z = \max_i \{w_i\}$ , then the secret  $s^{wz}$  is said to be the majority of secrets.

III. ALGORITHMS

Our aim is first to describe approach to detect and identify cheaters. Then, we propose our scheme which is based on Shamir's  $(t, n)$ -SS scheme. One unique feature of our proposed scheme is that we use the same share for secret reconstruction to detect and identify cheaters. Our scheme is an extension of Shamir's  $(t, n)$ -SS scheme.

– *Method for detecting cheaters* In Shamir's  $(t, n)$ -SS scheme, a  $t - 1$  degree interpolating polynomial can be uniquely reconstructed based on  $t$  shares. Thus, if there are more than  $t$  shares and there is no faked share, a consistent polynomial should be reconstructed for all combinations of  $t$  shares. Cheater detection is determined by detecting inconsistent polynomials (or secrets) among all reconstructed secrets. However, cheaters can collaborate to determine their faked shares to fool honest shareholders to believe that a faked secret is a real secret. In Sec. 5, we will discuss bounds of detect ability of our proposed detecting scheme under three attacks as presented in next section.

– *Method for identifying cheaters* When cheaters have been detected, there are inconsistent reconstructed polynomials (or secrets) for all combinations of  $t$  shares. Among all reconstructed secrets, if the legitimate secret is the majority of secrets as we have defined in Def. 2, we can use the *majority voting mechanism* to identify each faked share. We need to investigate conditions that the

Fig.

legitimate secret is the majority of secrets. In addition, we will discuss bounds of identifiability of our proposed identifying scheme under three attacks as presented in next section.

We use  $c$  to denote the number of faked shares and  $j$  ( $n \geq j \geq t$ ) to denote the number of participants in a secret reconstruction. There are  $j - c$  legitimate shares in a secret reconstruction.

**Algorithm 1 (Cheater detection)**

Input:  $t, n, J, si_1, \dots, si_j$

1. Compute an interpolated polynomial  $f(x)$  of  $j$  points  $(i_1, si_1), \dots, (i_j, si_j)$ . Set the degree of  $f(x)$  to be  $d$ .
2. If  $d = t - 1$ , then  $s = f(0)$ , and

Output: There is no cheater and Secret is  $s$ ; otherwise

Output: There are cheaters.

**Algorithm 2 (Cheater identification)**

Input:  $t, n, s, J, T, si_1, \dots, si_j$

1. For all  $Ti \in \tau$ , compute  $si = F(Ti)$  where  $i = 1, \dots, u$ .
2. Divide  $U = \{s^1, \dots, s^u\}$  into  $v$  subsets  $U_i$  such that  $U = U_1 \cup \dots \cup U_v$  where  $U_k \cap U_l = \emptyset$  for  $1 \leq k, l \leq v$  and  $k \neq l$ , and  $U_i = \{s^{i1}, \dots, s^{iwi}\}$  where  $s^{wi} = s^{i1} = \dots = s^{iwi}$ .
3. Set  $wz = \max_i \{w_i\}$ , and set  $s = s^{wz}$ .
4. Pick  $Tk \in T$  such that  $s = F(Tk) = F Tk (s_{ik1} \dots s_{ikt})$ , and set  $R = J - \{i_{k1}, \dots, i_{kt}\}$ .
5. Pick  $i_r \in R$  orderly and remove it from  $R$ , and compute  $s^r = F(s_{ir}, s_{ik2}, \dots, s_{ikt})$ .
6. If  $s^r = s$ , then put  $i_r$  into  $H$ ; otherwise put  $i_r$  into  $C$ .
7. Return Step 5 until  $R = \emptyset$ .

Output: The cheater set is  $C$ .

*Remark 2* The computational complexity of algorithm 1 is  $O(1)$  and the complexity of algorithm 2 is  $O(j!)$ , where  $j \leq n$ . We want to point out that  $n$  is the total number of shares in a secret sharing scheme and  $n$  is independent with the security of secret sharing scheme.

**IV .CHEATER ATTACK CLASSIFICATION**

Here, we discuss about three attacks of cheaters that are against our proposed detection and the identification scheme.

– *Type 1 attack* the cheaters of this type attack can be either honest shareholders who present their shares in error *accidentally* or dishonest shareholders who present their faked shares *without* any collaboration. Each faked share of this attack is just a random integer and is completely independent with other shares.

– *Type 2 attack* the cheaters of this type attack are dishonest shareholders who modify their shares on purpose to fool honest shareholders. In this type attack, we assume that all shareholders release their shares *synchronously*. Thus, cheaters can only collaborate among themselves to figure out their faked shares before secret reconstruction; but cannot modify their shares after knowing honest shareholders' shares (i.e. we assume that all shares must be revealed simultaneously). Under this assumption, only when the

number of cheaters is larger than or equal to the threshold value  $t$ , the cheaters can implement an attack successfully to fool honest shareholders.

– *Type 3 attack* the cheaters of this type attack are dishonest shareholders who modify their shares on purpose to fool honest shareholders. In this type attack, we assume that all shareholders release their shares *asynchronously*. Since shareholders release their shares one at a time, the optimum choice for cheaters is to release their shares after all honest shareholders releasing their shares. The cheaters can modify their shares accordingly. We consider the worst-case analysis to determine the bounds of detectability and identifiability of our proposed scheme.

**A. Cheater detection**

The cheater problem is a serious obstacle for secret sharing schemes. A cheater is a qualified participant who possesses a true share, but releases a fake share or withholds a share during a reconstruction of the secret. If a cheater releases a fake share or withholds a share on secret reconstruction, then he/she can obtain the secret and exclude others. Thus, the cheater has an advantage over the other shareholders.

**V. PROPOSED WORK**

Cheater identification scheme is based on Lagrange interpolation. In the  $(t,n)$ -threshold scheme proposed in this chapter a secret is an integer number  $S$ . Secret sharing schemes protect the secrecy and integrity of information by distributing the information over different locations. The  $(t, n)$  threshold secret sharing schemes were introduced by Shamir and Blakley independently in 1979 for protecting the cryptographic keys. Generation of shares and reconstruction of shares are challenging task in cheaters scenario. Cheaters identification is critical task on the time of share reconstruction. In this dissertation we proposed a roust secret share generation technique such technique based on cyclic point intersection of langrage's interpolation. In the process of share generation, construction and cheater identification, we proposed four steps. (i) Cyclic share generation (ii) share reconstruction and (iii) cheater identification. The proposed scheme used some notations are defined we assume that  $P$  is a participant set that contain  $n$  participant  $p_1, p_2, p_3, \dots, p_n$ . Such that  $p = \{p_1, p_2, p_3, \dots, p_n\}$  and  $c_1, c_2 \dots c_n$  are cyclic prefix of interpolation equation. Each member of  $P$  shares a secret  $K$  and hold a secret cyclic prefix  $C_i$  where  $1 \leq i \leq n$ .

**A. Share generation phase.**

Assume that a dealer wants to share a secret  $K$  among the  $n$  members in  $P$ . First, the dealer specifies the threshold value  $t$  freely within the range  $1 \leq t \leq n$ . then dealer select three point of prime in subsequent in cyclic  $x, y, z$ . The dealer randomly generates  $n$  different polynomials  $f_i$ 's of degree  $t-1$ , such that

$$F_i(X) = a(i, 0) + a(i, 1)X + \dots + a(i, t-1)X^{t-1}$$

Now then the cyclic point of intersection put into each generated shares  $X_c, Y_c$  and  $Z_c$

As

Consider two distinct points J and K such that  $J = (x_cJ, y_cJ)$  and  $K = (x_cK, y_cK)$

Let  $L = J + K$  where  $L = (x_cL, y_cL)$ , then

$$x_cL = s^2 - x_cJ - x_cK$$

$$y_cL = -y_cJ + s(x_cJ - x_cL)$$

$s = (y_cJ - y_cK)/(x_cJ - x_cK)$ ,  $s$  is the slope of the line through J and K.

If  $K = -J$  i.e.  $K = (x_cJ, -y_cJ)$  then  $J + K = O$ . where O is the point at infinity. If  $K = J$  then  $J + K = 2J$  then point

interpolation the difference value of cyclic prefix is 0 there is no cheater and the cyclic point generate a difference 1 then there is cheater.

VI.ACKNOWLEDGMENTS

I specially thanks my M.Tech guide Mrs.Megha Kamble who help me time to time.For this work my parents also insist me so I am also thankful of them

VII.CONCLUSIONS

In this paper, we consider the cases when there are more than  $t$  shareholders participated in secret reconstruction. Since there are more than  $t$  shares for reconstructing the secret, the redundant shares of a  $(t, n)$  secret sharing scheme can be used to detect and identify cheaters. We introduce the property of consistency and the notion of the majority of secrets to detect and identify cheaters. The bounds of detectability and identifiability under three attacks are presented. We utilizes shares for secret reconstruction to detect and identify cheaters. Our scheme is an extension of Shamir’s secret sharing scheme.

REFERENCES

- [1] W. Trappe and I.C.Washington, “Introduction to Cryptography with Coding Theory”, Pearson International Edition (2006)
- [2] C. C. Thien and J.C.Lin, “Secret image sharing,” Computers & Graphics, vol. 26, no.1, 765-770, 2002.
- [3] Blakley G.R., “Safeguarding cryptographic keys. In: Proceedings of AFIPS’79”, vol. 48, pp. 313-317 (1979).
- [4] Brickell E.F., Stinson D.R., “ The detection of cheaters in threshold schemes. In: Proceedings of Crypto’88”, LNCS, vol. 403, pp. 564-577. Springer-Verlag (1990).
- [5] CarpentieriM., ” A perfect threshold secret sharing scheme to Identify cheaters”, Des. Codes Cryptogr. 5(3), 183-187 (1995).
- [6] CarpentieriM., De Santis A., Vaccaro U, ” Size of shares And Probability of cheating in threshold”, schemes. In: Proceedings of Eurocrypt’93, LNCS, vol. 765, pp. 118 125. Springer-Verlag (1994).
- [7] Charnes C., Pieprzyk J., Safavi-Naini R., ”Conditionally Secure secret sharing scheme with disenrollment Capability” In: Proceedings of CCS’94, pp. 89-95. ACM (1994).
- [8] Kurosawa K., Obana S., Ogata W.:  $t$ -cheater identifiable  $(k, n)$  secret sharing schemes. In: Proceedings of Crypto’95, LNCS, vol. 963, pp. 410-423. Springer-Verlag (1995).
- [9] Rabin T., Ben- Or M., “Verifiable secret sharing and multiparty protocols with honest majority”, In:Proceedings of the 21st Annual ACM Symposium on the Theory of Computing, pp. 73-85 (1989).
- [10] Shamir A., “ How to share a secret. Comm.”, ACM 22(11), 612-613 (1979).
- [11] Simmons G., “An introduction to shared secret schemes and their applications”, Sandia Report SAND 88-2298 (1988).
- [12] Tompa M., Woll H, “ How to share a secret with cheaters”, J. Cryptol. 1(3), 133-138 (1989).

Schemes	Shamir	Blakley	Tompa	Wall	Proposed
Perfect sharing	Yes	No	No	No	Yes
Size of share	Same	Small	Larger	Larger	Same
Method used for SSS	Polynomial	Hyper plane	One way hash	One way hash	Polynomial
Reveal info	Yes	Yes	No	No	No
Cheater detection	No	No	Yes	Yes	Yes
Cheater identification	No	No	Yes	Yes	Yes

doubling equations are used.Then dealers send the all generated shares to participant.

B. The Secret Reconstruction Phase

Assume that the participants P1, P2. Pr of any qualified subset in P wants to Cooperate to reconstruct the shared secret K. They can perform the following steps To determine the shared secret K. In the reconstruction phase we apply cyclic addition point of interpolation.

Consider a point J such that  $J = (x_cJ, y_cJ)$ , where  $y_cJ \neq 0$

Let  $L = 2J$  where  $L = (x_cL, y_cL)$ , Then

$$x_cL = s^2 - 2x_cJ \text{ mod } p$$

$$y_cL = -y_cJ + s(x_cJ - x_cL) \text{ mod } Z_c$$

$s = (3x_cJ^2 + a) / (2y_cJ) \text{ mod } Z_c$ ,  $s$  is the tangent at point J and  $a$  is one of the parameters chosen with the elliptic curve. If  $y_cJ = 0$  then  $2J = O$ , where O is the point at infinity.

C. Cheaters detection phase

In the cheater detection phase ,reconstructed shares find the point of intersection of cyclic in language’s

# Architecture of GPRS and Its Capability Comparison with HSDPA

A. Sarthak R. Patel

**Abstract** - The aim of this paper is to understand the architecture of the GPRS and focus primarily on the issue come up while migrating from existing GSM network to that of GPRS. The migration path from GSM to GPRS requires additional packet switching nodes, software upgrades in the base station subsystem. High-Speed Downlink Packet Access (HSDPA) (Sometimes known as High-Speed Downlink Protocol Access) is a 3G mobile telephony protocol in the HSPA family, which provides a roadmap for UMTS-based networks to increase their data transfer speeds and capacity.

**Index Terms** – GPRS, GSM, HSDPA, UMTS, EDGE, W-CDMA

## I. INTRODUCTION

WIRELESS data represents an increasing percentage of operator revenues. Beginning with the success of Short Message Service (SMS) in Europe and iMode in Japan, users and enterprises are beginning to embrace wireless data in a wide range of other applications, including e-mail, game downloads, instant messaging, ringtones, video, and enterprise applications such as group collaboration enterprise resource planning. Enhanced Data Rates for GSM Evolution (EDGE) has proven to be a remarkably effective cellular-data technology, and is now supported by operators and vendors worldwide. Meanwhile, deployment of UMTS, which has even more powerful data capabilities, is accelerating. High Speed Downlink Packet Access (HSDPA), an enhancement for Universal Mobile Telecommunications System (UMTS), will soon increase data capabilities even further [1]. The result is a portfolio of complementary technologies that realize the potential of wireless data—the GSM family of data technologies.

Wireless data represents an increasing percentage of operator revenues. Beginning with the success of Short Message Service (SMS) in Europe and iMode in Japan, users and enterprises are beginning to embrace wireless data in a wide range of other applications, including e-mail, game downloads, instant messaging, ringtones, video, and enterprise applications such as group collaboration enterprise resource planning.

Though voice still constitutes most of cellular traffic and the bulk of service revenue, [2] wireless data is beginning to appreciably add to service revenue. In a recent quarter one GSM operator (MM02) reported that over 20 percent of their service revenue was from wireless data. There are a number of important factors that are accelerating adoption of wireless data, including increased user awareness, innovative devices such as smart phones and global coverage. But two factors stand out: network capability and applications. Technologies such as EDGE and UMTS provide the capability to support a wide range of applications, including standard networking applications as well as those designed for wireless. Meanwhile, application and content suppliers are optimizing their applications, or in many cases developing entirely new applications and content to target the needs and desires of mobile users.

Computing itself is becoming ever more mobile, and notebooks, PDAs and smart phones are now prevalent. In fact, all phones are becoming “smart” with some form of data capability. Lifestyles and work styles themselves are increasingly mobile, with more and more people traveling for work, pleasure, or in retirement. Meanwhile, the Internet is becoming progressively more intertwined in the fabric of people’s lives, providing communications, information, enhancements for memberships and subscriptions, community involvements, and commerce. In this environment, wireless access to the Internet is a powerful catalyst for the creation of new services and new business opportunities for operators as well as third-party businesses.

As the benefits of these services become apparent, as the services themselves become more powerful thanks to higher throughput rates and quality-of-service mechanisms, and as service costs drop due to increased spectral efficiency, use will constantly grow. Wire line data already represents more than fifty percent of traffic within worldwide telecom networks. Though cellular data represents just a small portion of service revenue today, in all likelihood, similar growth will happen with cellular networks.

With data constituting a rising percentage of total cellular traffic, it is essential that operators deploy data technologies that meet customer requirements for performance and that are spectrally efficient, especially as data applications can demand significant network resources. Operators have a huge investment in spectrum and in their networks—data services must leverage these investments. It is only a matter of time before today’s over 1.5 billion cellular customers start fully taking advantage of data capability. This presents tremendous

opportunities and risks to operators as they choose the most commercially viable evolution path for migrating their customers.

The GSM family of data technologies, including General Packet Radio Service (GPRS), EDGE, UMTS - Wideband CDMA (WCDMA) and HSDPA, provides a powerful set of capabilities, spectral efficiencies, and means of deployment that maximizes revenue and profit potential[7].

Some of the important observations and conclusions of this paper include:

- 1) Before EDGE was commercially deployed, a previous version of this paper projected performance gains for EDGE. Results from the field were exactly as predicted – EDGE more than triples GPRS data throughputs, delivering typical rates of 100 to 130 kilobits per second (kbps).
- 2) UMTS is spectrally extremely efficient for high-data throughput services. It also offers high peak rates of 2Mbps with average rates between 220 kbps and 320 kbps, multimedia support, (e.g., conversational video), quality of service, and a smooth upgrade for future enhancements.
- 3) HSDPA will be even more effective for enhancing WCDMA performance than EDGE was for enhancing GPRS performance, with a standard supporting peak rates of 14 Mbps and average throughput rates close to 1 Mbps.
- 4) Unlike some competing technologies, GSM/GPRS/EDGE/WCDMA/HSDPA allows an operator to efficiently use their entire available spectrum for voice and data services.
- 5) There is accelerating momentum globally in the deployment of EDGE and UMTS. GPRS/EDGE and UMTS are being deployed as complementary 3G technologies.
- 6) As one of the first cellular technologies to feature adaptive modulation and coding schemes and incremental redundancy, EDGE is spectrally very efficient for medium-bandwidth applications.
- 7) Operators will be able to do a simple software upgrade of their UMTS networks to support HSDPA much as they did to their GPRS networks to support EDGE.
- 8) With a UMTS multi-radio network, a common core network can efficiently support GSM, GPRS, EDGE, WCDMA, and HSDPA access networks, offering high efficiency for both high and low data rates, and for high and low traffic density configurations.
- 9) Ongoing UMTS evolution includes significant enhancements with each new specification release, including higher throughput rates, enhanced multimedia support, and integration with wireless LAN technology.

This paper begins with an overview of the market, looking at adoption of services, deployment of GSM-UMTS technologies and other wide-area wireless technologies. It then explains the capabilities as well as workings of the different technologies, including GPRS, EDGE, WCDMA and HSDPA[8]. The emphasis of this section is to quantify real-world performance. The paper then examines the evolution from GPRS to HSDPA, including how increasing spectral efficiency will drive deployment.

## II. MARKET

In considering the market, we review market adoption of wireless data[2], deployments of GSM/UMTS networks around the world, and how GSM/UMTS relates to other wireless technologies.

### A. Market Adoption

While wireless data has always offered a tantalizing vision of always-connected mobile computing, adoption has been slower than for voice services. In the past two years, however, adoption has accelerated, finally some might say, thanks to a number of key developments. Networks themselves are much more capable, delivering higher throughputs at lower cost. Awareness of data capabilities has increased, especially through the widespread success of SMS, wireless e-mail, downloadable ringtones and downloadable games. Widespread availability of services has also been important. The features found in cellular telephones is expanding at a rapid rate, and today includes large colour displays, graphics viewers, still cameras, movie cameras, MP3 players, instant messaging clients, e-mail clients, push-to-talk, downloadable executable content capability, and browsers supporting multiple formats. All of these capabilities consume data. Meanwhile, smart phones with their emphasis on a rich computing environment on the phone; represent the convergence of the personal digital assistant, a fully capable mobile computer and a phone in a device that is only slightly larger than the average cellular telephone. Many users would prefer to carry one device that “does it all.”

As a consequence, this rich network and device environment is spawning the availability of a large range of wireless applications and content. Why? Application and content developers simply cannot afford to ignore this market because of its growing size, and its unassailable potential. And they aren't. Consumer content developers are already successful by providing downloadable ringtones and games. [6]Enabled by 3G network capability, downloadable and streaming music and video are not far behind. In the enterprise space, all the major developers now offer mobilized “wireless-friendly” components for their applications. Acting as catalysts, a wide array of middleware providers address issues such as increased security (e.g., VPNs), switching between different networks (e.g., WLAN to 3G) and session maintenance under adverse radio conditions.

The wireless-data market has not yet reached the critical mass where adoption and content explode, but it is close, as suggested by the following market data. In 2003, wireless Internet usage grew 145% with 134 million people trying or using services. This figure could grow to 600 million users by 2008. Mobile Java applications are expected to grow from \$1.4 billion in 2003 to \$15.5 billion in 2008.

According to a recent study by In-Stat MDR released July 15, 2004 “Understanding Decision-makers’ and Decision-influencers’ Dual Roles in the Implementation of Wireless Data in the Business Environment”,

- 1) Wireless data is becoming more important to enterprise companies. About 20% of IT budgets are dedicated to

wireless data, which was defined as including wireless LAN and WAN equipment and services.

- 2) About half of mid-size and large enterprises use wireless data today and another 30 percent are planning/evaluating future use. Three quarters of current WAN data user plan to increase their usage in the future.
- 3) Common wireless data applications include email, access to the Web, and to spreadsheets and word processing documents. Future applications include instant messaging and Web-based applications.

This market data is encouraging, but realistically, the market is still in its infancy. Though awareness of services is higher than ever before, many people still do not understand the true range of data options available to them. For example, how many business users realize they can use their Bluetooth-equipped phone as a modem for their laptops? The number of enhanced mobile data applications is still low relative to its market potential. For example, it should be possible to request a taxi with one simple request on a mobile telephone which notifies the taxi company of a user’s exact location, dispatches a taxi, and sends update messages indicating the taxi’s arrival time. Services like this are coming, but they are not available yet as they require the integration of existing dispatch systems with geographic databases, location-based services and mobile commerce systems. In the enterprise space, the first stage of wireless data was essentially to replace modem connectivity. The next is to offer existing applications on new platforms such as smart phones. But the final, and much more important change, is where jobs are reengineered to take full advantage of continuous connectivity. All this takes time, but the momentum in the direction of increased efficiency, increased convenience and increased entertainment, all fuelled by wireless data, is unstoppable.

The key for operators is to offer networks that can support the demands of wireless consumer and business applications as they grow, as well as to offer the complementary capabilities, such as message stores, e-commerce solutions, location-based services, QoS. This is where the GPRS to HSDPA data story is particularly compelling, because not only does it provide a platform for continual improvements in capabilities, but it does so over huge coverage areas and on a global basis.

**B. Deployments**

The size of the potential data market is extensive, and can readily be appreciated by looking at the scope of GSM/UMTS deployments that are occurring. Today, more than one billion subscribers are using GSM. Nearly every GSM network in the world today supports GPRS[3].

EDGE is another success story. As of July, 2004, 114 operators in 68 countries from all over the world were working with EDGE. This includes 34 operators offering commercial service and 38 operators in active deployment. EDGE has reached critical mass in terms of POPS, geography, infrastructure, and devices. EDGE operators represent over half a billion potential EDGE customers within their networks. Due to the very small incremental cost of including EDGE capability in GSM network deployment,

virtually all new GSM infrastructure deployments are likely to be EDGE capable with virtually all new mid to high level GSM devices including EDGE as well.

[3]UMTS deployments are also accelerating. There are forty-six commercial UMTS networks already in operation in twenty-four countries, with seventy-one more either pre-commercial, in planning, licensed or currently being deployed.

The Shosteck Group anticipates that, “During 2007, we estimate 70 million new subscribers, bringing the total to 125-150 million...” In addition, the firm predicts that 110 million UMTS handsets will be sold in 2007. Other leading analyst firms have predicted even higher subscriber counts.

**III. GPRS**

GPRS is the world’s most ubiquitous wireless data service, available in over 201 countries, with service from 217 operators and a choice of over 591 handsets. Various analysts predict unit sales of over 150 million GSM/GPRS devices in 2004. GPRS is a packet-based IP connectivity solution supporting a wide range of enterprise and consumer applications. GPRS networks operate as wireless extensions to the Internet, and give users Internet access as well as access to their organizations from anywhere. With average throughput rates of up to 40 using four time-slot devices, users have the same effective access speed as a modem, but with the convenience of being able to connect from anywhere.

To understand the evolution of data capability, we examine briefly how these data services operate, beginning first with the architecture of GPRS, as depicted in Fig. 1.

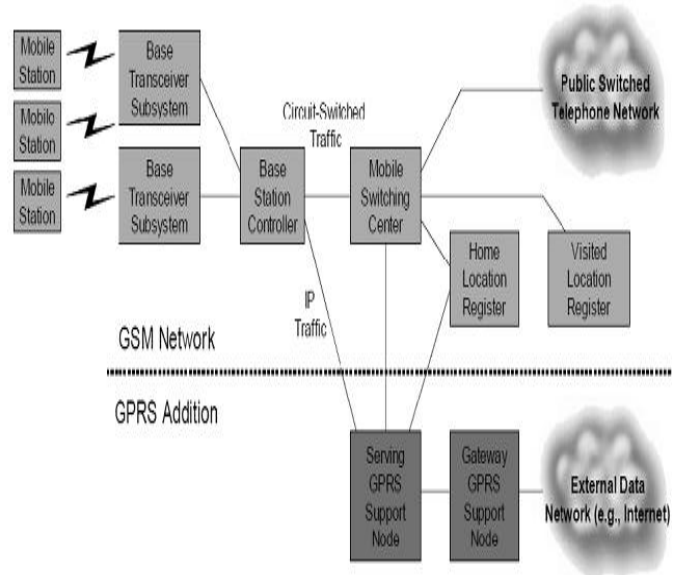


Fig. 1 GPRS Architecture

GPRS is essentially the addition of a packet-data infrastructure to GSM. The functions of the data elements are as follows[5]:

- 1) The base station controller directs packet data to the Serving GPRS Support Node (SGSN), an element that authenticates and tracks the location of mobile stations.

- 2) The SGSN performs the types of functions for data that the mobile switching centre performs for voice. There is one SGSN for each serving area, and it is often collocated with the MSC.
- 3) The SGSN forwards user data to the Gateway GPRS Support Node (GGSN), which is a gateway to external networks. There is typically one GGSN per external network (e.g., Internet). The GGSN also manages IP addresses, assigning IP addresses dynamically to mobile stations for their data sessions.

Another important element is the home location register (HLR), which stores users' account information for both voice and data service. What is significant is that this same data architecture supports data services in EDGE and UMTS networks, simplifying operator network upgrades.

In the radio link, GSM uses radio channels of 200 KHz width, divided in time into eight time slots that repeat every 4.6 msec. The network can have multiple radio channels (referred to as transceivers) operating in each cell sector. The network assigns different functions to each time slot, such as the broadcast control channel, circuit switched functions like voice calls or circuit-switched data calls; the packet broadcast control channel (optional), and packet data channels. The network can dynamically adjust capacity between voice and data functions, and can also reserve a minimum amount of resources for each service. This enables more data traffic when voice traffic is low or likewise more voice traffic when data traffic is low, and maximizes the overall use of the network.

GPRS offers close coupling between voice and data services. While in a data session, users can accept an incoming voice call, which suspends the data session, and then resume their data session automatically when the data session ends. Users can also receive SMS messages and data notifications while on a voice call. EDGE networks and devices behave in the same way.

With respect to data performance, each data time slot can deliver user data rates of about 10 kbps using coding schemes 1 and 2, and the network can aggregate up to four of these on the downlink with current devices to deliver users perceived data throughputs of up to 40 kbps. If there are multiple data users active in a cell sector, they share the available data channels. However, as demand for data services increases, operators can accommodate customers by assigning an increasing number of channels for data service limited only by their total available spectrum and radio planning.

With coding schemes 3 and 4, GPRS has greater flexibility in how the radio link allocates communicated bits between data and error control, resulting in increased throughput with higher signal quality. The result is throughput rates up to 33% higher and increased overall spectral efficiency of about 30%. Coding schemes 3 and 4 are an option for operators. To boost GPRS performance and capacity even further, operators are deploying EDGE technology.

#### IV. HSDPA

High Speed Downlink Packet Access is a tremendous performance upgrade to WCDMA for packet data that delivers peak rates of 14 Mbps and that is likely to increase

average throughput rates to about 1 Mbps, a factor of up to three and a half times over WCDMA[4]. HSDPA also increases spectral efficiency by a similar factor. Available in 3GPP Rel'5, operators will trial HSDPA in 2004 and 2005 with commercial availability by the end of 2005 or early 2006. NTT DoCoMo and Cingular Wireless, with HSDPA planned for launch in 2005, are expected to be the first operators to deploy HSDPA. HSDPA is fully backwards compatible with WCDMA, and any application developed for WCDMA will work with HSDPA. The same radio channel can simultaneously service WCDMA voice and data users, as well as HSDPA data users. HSDPA will also have significantly lower latency, expected at close to 100 msec.

HSDPA achieves its high speeds through similar techniques that amplify EDGE performance past GPRS, including higher-order modulation, variable coding and incremental redundancy, as well as through the addition of powerful new techniques such as fast scheduling. [8]HSDPA takes WCDMA to its fullest potential for providing broadband services, and is the highest-throughput cellular-data capability defined. The higher spectral efficiency and higher speeds not only enable new classes of applications, but also support a greater number of users accessing the network.

HSDPA achieves its performance gains from the following radio features:

- 1) High speed channels shared both in the code and time domains.
- 2) Short transmission time interval (TTI).
- 3) Fast scheduling.
- 4) Higher-order modulation.
- 5) Fast link adaptation.
- 6) Fast hybrid automatic-repeat-request (ARQ).

These features function as follows. First, HSDPA uses high speed data channels called High Speed - Downlink Shared Channels (HS-DSCH). Up to 15 of these can operate in the 5 MHz WCDMA radio channel. Each uses a fixed spreading factor of 16. User transmissions are assigned to one or more of these channels for a short transmission time interval of 2 msec, significantly less than the interval of 10 to 20 msec used in WCDMA. The network can then readjust which users are assigned which HS-DSCH every 2 msec. The result is that resources are assigned in both time (the TTI interval) and code domains (the HS-DSCH channels).

Fast scheduling exploits the short TTI by assigning channels to the users with the best instantaneous channel conditions, rather than in a round-robin fashion. Since channel conditions vary somewhat randomly across users, most users can be serviced using optimum radio conditions, and can hence obtain optimum data throughput. The system also makes sure that each user receives a minimum level of throughput. The result is referred to as "proportional fair scheduling." Fig. 2 illustrates different users obtaining different radio resources.



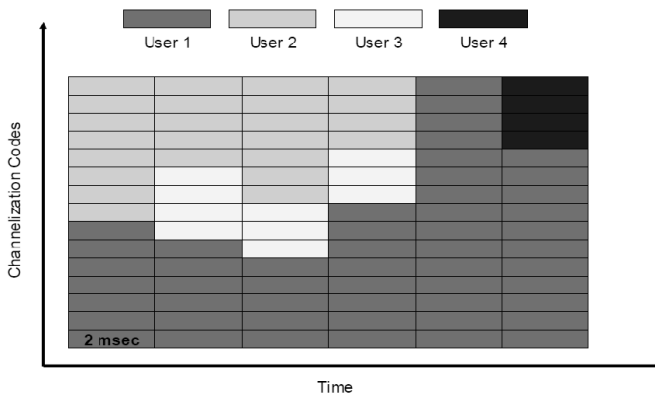


Fig. 2 High Speed – Downlink Shared Channels

HSDPA uses both the modulation used in WCDMA, namely Quadrature Phase Shift Keying (QPSK) and under good radio conditions, an advanced modulation scheme, 16 Quadrature Amplitude Modulation (16 QAM). The benefit of 16 QAM is that four bits of data are transmitted in each radio symbol as opposed to two with QPSK. 16 QAM increases data throughput, while QPSK is available under adverse conditions.

Depending on the condition of the radio channel, different levels of forward error correction (channel coding) can also be employed. For example, a three quarter coding rate means that three quarters of the bits transmitted are user bits and one quarter are errors correcting bits. The process of selecting and quickly updating the optimum modulation and coding rate is referred to as fast link adaptation. This is done in close coordination with fast scheduling described above.

Table 1 shows the different throughput rates achieved based on the modulation, the coding rate, and the number of HS-DSCH codes in use. Note that the peak rate of 14.4 Mbps occurs with a coding rate of 4/4, 16 QAM and all 15 codes in use.

TABLE 1 HSDPA THROUGHPUT RATES

Modulation	Coding Rate	Throughput with 5 codes	Throughput with 10 codes	Throughput with 15 codes
QPSK	1/4	600 Kbps	1.2 Mbps	1.8 Mbps
	2/4	1.2 Mbps	2.4 Mbps	3.6 Mbps
	3/4	1.8 Mbps	3.6 Mbps	5.4 Mbps
16 QAM	2/4	2.4 Mbps	4.8 Mbps	7.2 Mbps
	3/4	3.6 Mbps	7.2 Mbps	10.7 Mbps
	4/4	4.8 Mbps	8.4 Mbps	14.4 Mbps

Another HSDPA technique is referred to as Fast Hybrid Automatic Repeat Request (Fast Hybrid ARQ.) “Hybrid” refers to a process of combining repeated data transmissions with prior transmissions to increase the likelihood of successful decoding and “fast” refers to the error correcting mechanisms being implemented in the Node-B (along with scheduling and link adaptation), as opposed to the Base Station Controller in GPRS/EDGE. Managing and responding to real-time radio variations at the base station as opposed to an internal network node reduces delays and further improves overall data throughput.

Using the approaches just described, HSDPA maximizes data throughputs, maximizes capacity and minimizes delays.[8] For users, this translates to better network performance under loaded conditions, faster application

performance, a greater range of applications that function well, and increased productivity.

Despite HSDPA’s capabilities, researchers and developers are working on additional enhancements. First devices will support five codes with a peak rate of 3.6 Mbps. Subsequent devices will support ten to fifteen codes with a peak rate of 10.7 Mbps.

Other enhancements include two-branch diversity reception and equalizers in mobile devices. These improvements will occur one to two years after the initial deployment of HSDPA. Simulations show these features to further improve user data rates and network capacity. Relative to WCDMA Rel’99, these features will increase HSDPA performance from a factor of 2.5 to a factor of 3.5. Further evolution of HSDPA peak data rates can be achieved with multiple-input multiple-output (MIMO) antenna techniques of 3GPP Rel.’6.

### V. THE PATH FROM GPRS TO HSDPA

GPRS to HSDPA offers an increasing range of capabilities, supporting ever more demanding applications. GPRS, now available globally, already makes a wealth of applications feasible, including enterprise applications, messaging, e-mail, Web browsing, consumer applications, and even some multimedia applications. EDGE significantly expands the capability of GPRS, enabling richer Internet browsing, streaming applications, a greater scope of enterprise applications, and more multimedia applications. Then with UMTS and HSDPA[8], users can look forward to video phones, high-fidelity music, rich multimedia applications, and efficient access to their enterprise applications.

It is important to understand just what is important for users of these services, whether consumers or enterprises. The obvious needs are broad coverage, high data throughput and for enterprises, security. Less obvious needs, but as critical for effective application performance, is low latency, quality-of-service (QoS) control and spectral efficiency. Spectral efficiency, in particular, is of paramount concern, as it translates to higher average throughputs (and thus more responsive applications) for more users active in a coverage area. The discussion below, which examines each technology individually, shows how the progression from GPRS to HSDPA is one of increased throughput, increased security, reduced latency, improved QoS and increased spectral efficiency.

It is also helpful to specifically note the throughput requirements required for different applications. They are the following:

- 1) Micro browsing (e.g., WAP): 8 to 16 kbps.
- 2) Multimedia messaging: 8 to 32 kbps.
- 3) Video telephony: 64-384 kbps.
- 4) General purpose web browsing: 32 kbps to 384 kbps.
- 5) Enterprise applications, including e-mail, database access, virtual private networking: 32 kbps to 384 kbps.
- 6) Video and audio streaming: 32-384 kbps

This section discusses the evolution of data capability from GPRS to HSDPA and the stages available to operators to evolve their networks.[1] This progression, as shown in Fig. 3, happens in multiple phases, first with GPRS, then EDGE, then WCDMA, followed by evolved 3G capabilities such as

HSPDA, the Internet Multimedia Subsystem and eventually all-IP networks.

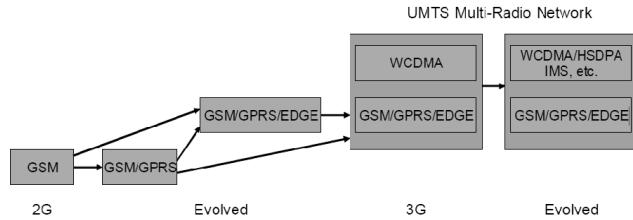


Fig. 3 Evolution of Cellular Technologies

GSM operators first enhanced their networks to support data capability through the addition of GPRS infrastructure, with the ability to use existing cell sites, transceivers and interconnection facilities. Operators who deployed GSM more recently (e.g., AT&T Wireless, Cingular Wireless, Rogers Wireless, Telecom Personal) installed GSM and GPRS simultaneously. More recently, operators have been upgrading their GPRS networks to EDGE with extremely good results.

## VI. THE FUTURE

HSDPA is a critical first step in enhancing W-CDMA network performance. However, there are already plans afoot to push the technology further. The next two major features of 3GPP's Release 6 and 7 specifications are an up-link improvement solution called enhanced up-link dedicated channel (EU-DCH) and throughput increases through multiple-input, multiple-output (MIMO).

While HSDPA is about maximizing downlink performance, EU-DCH is being developed to improve uplink-intensive services or applications requiring a balanced link like interactive video or voice over IP. EU-DCH will also reduce latency and facilitate the efficient use of TCP/IP applications like Microsoft Outlook over mobile networks.

## VII. CONCLUSION

This paper has described the data capabilities of GPRS to HSDPA. This evolution occurs in successive stages, with each stage increasing data throughput, increasing spectral efficiency, reducing network latency and adding new features such as quality-of-service and multimedia support. The migration and benefits of the evolution from GPRS to HSDPA is both practical and inevitable. Combined with the ability to roam globally, huge economies of scale, widespread acceptance by operators, complementary services such as multimedia messaging and a wide variety of competitive handsets, the result is a compelling technology family for both users and operators. UMTS has already been selected by some one hundred operators and has support from nearly all major regional standardization bodies. It offers an excellent migration path for GSM operators and as well as an effective technology solution for Greenfield operators.

UMTS is further enhanced by the deployment of High Speed Downlink Packet Access, an extremely fast data service with anticipated average speeds of about 1 Mbps, and peak speeds of up to 14 Mbps, the highest rate available for any cellular technology. HSDPA achieves its high speeds through similar techniques that propel EDGE performance

past GPRS as well as through the addition of powerful new techniques such as fast scheduling. Like EDGE, HSDPA can be deployed as software-only upgrade and is currently being trialed with commercial availability expected in late 2005 or 2006.

With the continued growth in mobile computing, powerful new handheld computing platforms, an increasing amount of mobile content, multimedia messaging, mobile commerce, and location services, wireless data will inevitably become a huge industry. GPRS to HSDPA provides the most robust portfolio of technologies and the optimum framework for realizing this potential.

## REFERENCES

- [1] 3G Americas, "The Evolution of UMTS, 3GPP Release 5 and Beyond", June 2004.
- [2] AT&T Wireless, research data, July 2002, November 2003, June 2004, submission to 3G Americas.
- [3] Cingular Wireless, "Spectrum Efficiency Comparison, GSM vs. UMTS vs. 1XRTT" - research material, March 14, 2002, submission to 3G Americas.
- [4] Ericsson, White Paper, "GSM to WCDMA the Global Choice", 2002.
- [5] Ericsson, Nokia, Siemens. "Agreed Data Performance Characterization for EGPRS, WCDMA and CDMA2000 1XRTT", September 20, 2002, submission to 3G Americas.
- [6] Nortel, "What 3G Applications & Services to Launch?", June 2004, submission to 3G Americas.
- [7] Lucent, "UMTS Data Performance from Simulations and Field Data Measurements", submission to 3G Americas.
- [8] Motorola, "Performance of 3GPP High Speed Downlink Packet Access (HSDPA)", Robert Love, Amitava Ghosh, Weimin Xiao and Rapeepat Ratasuk, 2004.

# The Most Promising Femtocell Technology and Mobility Related Issues

A. Prof. Vachik S. Dave, B. Mr. Sankit R. Kassa

**Abstract**—Femtocells are low-power, low cost, backward compatible, user-installed and portable wireless access points that operate in licensed spectrum to connect standard mobile devices to a mobile operator’s network using residential DSL or cable broadband connections. There are many technical challenges for femtocells but Mobility Management is the main challenge for it to be considered. This paper analyzes the technical issues and proposes an efficient mobility management mechanisms and methods for NGWN (Next Generation Wireless Networks) femtocells. The two types of management of location management and handoff management must be considered in femtocells. Here the topics related to mobility management, femtocell characterization, access control, paging procedure and cell handover are discussed and tried to find out the solutions related to mobility management. The proposed mechanisms can be taken into consideration in the near future when the Next Generation Wireless Technology will come into existence.

**Index Terms**— Femtocells, Mobility Management, Access control, paging procedure, Types, Challenges and their solutions.

## I. INTRODUCTION

IN telecommunications, a femtocell is a small cellular base station, typically designed for use in a home or small business [9]. It connects to the service provider’s network via broadband (such as DSL or cable); current designs typically support 2 to 4 active mobile phones in a residential setting, and 8 to 16 active mobile phones in enterprise settings. A femtocell allows service providers to extend service coverage indoors, especially where access would otherwise be limited or unavailable. Although much attention is focused on WCDMA, and the concept is applicable to all next generation wireless networks and standards, including GSM, CDMA2000, TD-SCDMA, WiMAX and LTE solutions.

Regardless of the network, one of the most important and challenging problems for tether less communication and computing are mobility management. Mobility management enables telecommunication networks to locate roaming terminals for call delivery and to maintain connections as the terminal is moving into a new service area.

Thus, mobility management supports mobile terminals,

A. AP, Computer Engineering Department, Gandhinagar Institute of Technology, Gujarat, India. (e-mail: vachik.daver@ git.org.in).

B. National Institute of Technology, Hamirpur, Himachal Pradesh, India (e-mail: kassasankit@gmail.com).

allowing users to roam while simultaneously offering them incoming calls and supporting calls in progress.

## II. MOBILITY MANAGEMENT

In general, Mobility management contains two components: (1) location management and (2) handoff management.

### A. Location Management

Location Management is a two-stage process that enables the network to discover the current attachment point of the mobile user for call delivery, as shown in Fig. 1. The first stage is location registration (or location update). In this stage, the mobile terminal periodically notifies the network of its new access point, allowing the network to authenticate the user and revise the user’s location profile. The second stage is call delivery. Here the network is queried for the user location profile and the current position of the mobile host is found. Current techniques for location management involve database architecture design and the transmission of signaling messages between various components of a signaling network. As the number of mobile subscribers increases, new or improved schemes are needed to support effectively a continuously increasing subscriber population [1].

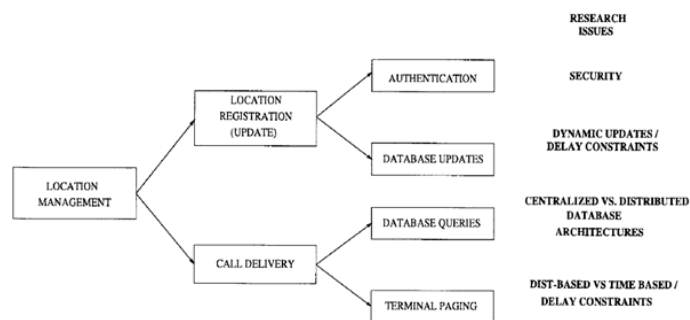


Fig 1. Location Management Operation

### B. Handoff Management

Handoff (or handover) management enables the network to maintain a user’s connection as the mobile terminal continues to move and change its access point to the network. The three-stage process for handoff first involves initiation, where the user, a network agent, or changing network conditions identify the need for handoff. The second stage is new connection generation, where the network must find new resources for the handoff connection and perform any additional routing operations. Under network-controlled handoff (NCHO), or

mobile-assisted handoff (MAHO), the network generates a new connection, finding new resources for the handoff and performing any additional routing operations. For mobile-controlled handoff (MCHO), the mobile terminal finds the new resources and the network approves. The final stage is data-flow control, where the delivery of the data from the old connection path to the new connection path is maintained according to agreed-upon service guarantees. The handoff management operations are presented in Fig. 2[1].

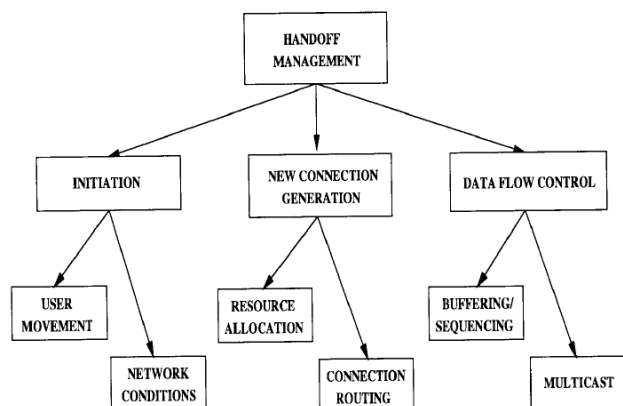


Fig 2. Handoff Management Operation

Handoff management includes two conditions: intracell handoff and intercell handoff. Intracell handoff occurs when the user moves within a service area (or cell) and experiences signal strength deterioration below a certain threshold that results in the transfer of the user's calls to new radio channels of appropriate strength at the same base station (BS). Intercell handoff occurs when the user moves into an adjacent cell and all of the terminal's connections must be transferred to a new BS. While performing handoff, the terminal may connect to multiple BS's simultaneously and use some form of signaling diversity to combine the multiple signals. This is called soft handoff. On the other hand, if the terminal stays connected to only one BS at a time, clearing the connection with the former BS immediately before or after establishing a connection with the target BS, then the process is referred to as hard handoff [8]. Handoff management research concerns issues such as: efficient and expedient packet processing; minimizing the signaling load on the network; optimizing the route for each connection; efficient bandwidth reassignment; evaluating existing methods for standardization; and refining quality of service for wireless connections. Fig. 2 lists these issues for the handoff management operations.

### III. FEMTOCELL MANAGEMENT

Femtocells are designed to produce extended coverage and enhanced capacity, especially in the indoor environment. Apparently, no specific Mobility Management (MM) for femtocells is necessary since they are expected to work within the existing network standards. However, the real implementation breaks many aspects of the current network assumptions. This is mainly due to the following factors:

1. Large number and high density of femtocells
2. Dynamic neighbor cell lists
3. Variant access methods
4. User/operator preference

#### A. Mobility Management of Femtocell in 3GPP

In the past few years, all of the femtocell vendors have made their own efforts with different structures and methods to fix the femtocell into the current network (UMTS, CDMA) or Wireless Interoperability for Microwave Access (WiMAX), etc.). Continuing on this path would result in over-hyped technology solutions that would make it difficult to inter-operate with each other and keep the cost of femtocells at a reasonable level. An industry wide standardization becomes essential to enable the widespread adoption and deployment of femtocells by telecom operators around the world. 3GPP started to work on a femtocell standard (including technical report (TR) and technical specification (TS)) in early 2008. Up-to-now, quite a number of proposals on femtocell architecture, mobility management and network security have been proposed.

With the freezing of release 8, some basic functionality in supporting the MM for femtocells has been achieved. However, to meet the restricted time line of release 8, only basic functionalities of MM for femtocells with closed access (known as CSG cells) are considered in release 8. Things like handover from macrocell to femtocell, handover between femtocells and supporting open and hybrid access method are still ongoing and will be handled in release 9.

3GPP has different Working Groups (WGs) working on different aspects of the radio interface and network architecture. The Technical Specification Group (TSG) Radio Access Network (RAN) is where all the radio air interface specifications are handled and contains six WGs as described in the following list:

1. RAN WG1 (RAN1) Radio layer 1 (Physical layer) specification,
2. RAN WG2 (RAN2) Radio layer 2 (MAC, RLC, PDCP) and Radio layer 3 RR,
3. RAN WG3 (RAN3) Iu, Iub, Iur, S1, X2 and UTRAN/E-UTRAN architecture,
4. RAN WG4 (RAN4) Radio performance and protocol aspects RF parameters and BS conformance,
5. RAN WG5 (RAN5) Mobile terminal conformance testing,
6. RAN AHG1 (RAN\_AH1) Ad-hoc group coordinating communications between 3GPP and ITU-R.

There are a couple of specifications related to MM available in both release 8 and release 9. Following are UMTS/LTE standards and their corresponding WGs, which involve the mobility procedures with Home NodeB (HNB)/Home eNodeB (HeNB).

Femtocells are quite new in the 3GPP standardization process, although many mobility procedures for femtocells have been addressed. Things like how these procedures actually work and their criteria are still under discussion and

further studies are needed. One way to find a more detailed and ongoing work on the specific methodology or parameter in the specification process is to go through the contribution papers (called technical documents, or tdocs) that members of 3GPP upload on their meeting website. Most of the MM procedures (access control, cell selection/reselection and handover, etc.) are discussed in RAN2 meetings. Paging procedure and access control are also discussed in RAN3. After TSG-RAN Meeting, the following mobility procedures are also within the scope of RAN3.

- Enhanced handover scenario:
- In-bound mobility 3G macro to HNB handover including legacy UE aspects,
- 3G HNB to 3G HNB handover,
- In-bound mobility LTE macro to HeNB handover,
- LTE HeNB to LTE HeNB handover.
- Enhanced access scenarios:
- Open access,
- Hybrid access.

#### IV. FEMTOCELL CHARACTERIZATION

It is very difficult or even infeasible to treat femtocells as normal macrocells. The network cannot afford broadcasting the femtocell information over the network, and the signaling overhead between the macrocell users and femtocells will have a big impact on the performance of the network. The identifiers and mechanisms that can characterize the aspects of the femtocells are necessary in reducing the impact and enhancing the mobility procedures of supporting femtocells.

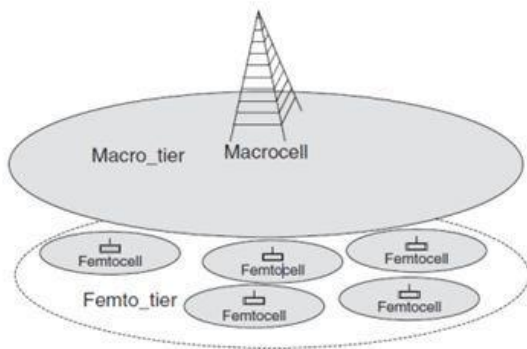


Fig. 3 Two layer network

##### A. Distinguish Femtocell from Macrocell

With the help of the identifiers and mechanisms that can distinguish femtocells from macrocells, the whole network can be treated as a two-layer network: macro-network and femto-network as shown in Figure 3. A few efforts can then be made to enhance the mobility procedures and reduce the signaling overhead between these two layers. For example, new reselection parameters can be configured between these two layers to prioritize the femtocell users to camp on once they are entering the coverage of the allowed femtocell. Also, these

parameters help in preventing the users that are not allowed to use femtocell from scanning or reading the information from the femtocell layer, leading to a longer battery life.

##### B. Separate Femtocell PLMN ID

Another method of distinguishing femtocells from macrocells is to have a separate femtocell PLMN Identity (PLMN ID) that is distinct from that of the macrocells [2]. In this approach, femtocells will be assigned a different PLMN ID from macrocells to secure femtocell selection and minimize impact on macrocell users as shown in Figure 4. UEs that are not allowed to use a femtocell could be configured not to access the femtocell Public Land Mobile Network (PLMN), resulting in better battery performance, and less signaling load

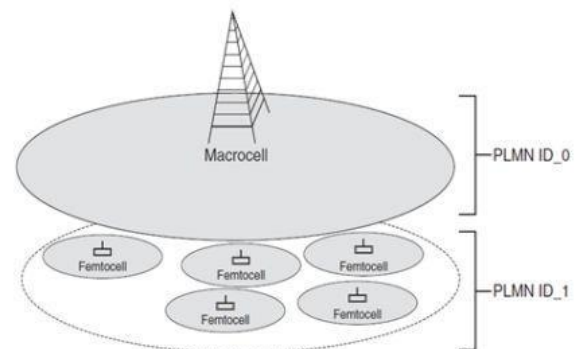


Fig 4. Separate femtocell PLMN Architecture

towards the Core Network (CN). A separate PLMN ID allows the UE to display the right network identifier, indicating to the user that he has camped on a femtocell. As the straightforward approach for the pre-release 8 network, this method is currently used by many of the femtocell vendors as the main method of distinguishing femtocells from macrocells. However, when using separate femtocell PLMN, additional PLMN IDs are required, which some of the network operators may not have. Also, some old Subscriber Identity Module (SIM)/Universal Integrated Circuit Card (UICC) card may have some compatibility problems with displaying the right femtocell PLMN name.

#### V. PROCESS OF FINDING NEIGHBORING FEMTOCELLS

By above two methods, it is easy to distinguish femtocells from macrocells. During inter-cell mobility it is very easy for the UE to know the surrounding macrocells using the NCL (Neighbour Cell List). But it is a real problem to find out the neighboring femtocells by UE because femtocells can be power ON or OFF anytime and this kind of process can cause a problem.

There are two types of efforts made to enable UE to find femtocells as below.

##### A. Neighbour cell list (NCL):

Two types of mobility are in NCL: Outbound Mobility and Inbound Mobility In outbound mobility, from a femtocell to macrocell, the NCL in the femtocell is a straight forward way

for the UE to be aware of the surrounding macrocells and femtocells.

The NCL can be created by the femtocell through self-configuration algorithms implemented by the vendors. The NCL may be updated each time the femtocell is powered on or when the femtocell senses any change of the neighbouring cells. For inbound mobility from macrocells, it is infeasible to include all of the neighbouring femtocells in the list due to the high volume of femtocells. However, alternative ways have been made by some vendors to support high density femtocells in pre-release 8 network. The main idea is based on the reuse of a certain number of PSCs (i.e. 10).

These PSCs can then be programmed into the macrocell's neighbour cell list. When such an NCL is received from the serving macrocell, the UE can then do measurements for these femtocells corresponding to the PSCs as for macrocells. The drawback is that the NCL of macrocells needs to be updated and PSC confusion may happen when the UE detects two nearby femtocells with the same PSC [3].

### B. UE Autonomous Search:

In 3GPP release 8, UE autonomous search function is introduced for the UE to find the CSG cells. The UE is required to perform an autonomous search function in order to detect suitable CSG cells [4, 5]. The UE, which cannot access CSG cells, can disable the autonomous search function for CSG cells. If 'Dedicated CSG frequency (ies)' information element (IE) is present, the UE may use the autonomous search function only on these dedicated frequencies and on the other frequencies listed in the system information. To assist the search function on shared carriers, UE may search cells with reserved PSCs/PCIs defined by CSG PSC/PCI Split Information for intra-frequency and inter-frequency measurements and mobility purposes.

## VI. ACCESS CONTROL

In a macro cellular network, Access control normally happens during UE requests a data transmission service. However, access control needs to be invoked whenever a UE is trying to camp on a femtocell to prevent the unauthorized use of that femtocell. Here access control for different access types is shown.

### A. Closed Access

Closed access mode femtocell is known as the CSG cell in 3GPP and is the only featured access mode in release 8. The CSG cell may be widely used in individual home deployment.

In this scenario, the owner of the femtocell does not want to share the femtocell due to the limited source of the backhaul or due to some security concerns. The access control should always be performed whenever a UE is trying to camp on the femtocell. Any UE that is not in the CSG will be rejected by the femtocell.

### B. Open Access

As a new access mode in 3GPP release 9, the open access

mode femtocell operates as a normal cell, i.e. non-CSG cell. The operator may deploy an open access mode femtocell to fill some indoor blind spots and some public hot-spots to serve all the users as a macrocell. The mobile network doesn't need to perform any specific UE access control for such a femtocell.

### C. Hybrid Access

This access feature will be available in 3GPP release 9. A hybrid mode means that the femtocell can provide a combination of both open and closed access modes at the same time. The hybrid access femtocell is a cell that not only has a CSG ID, but also allows UEs that are not members of that CSG to camp thereon. In this access mode, these UEs may only be authorized a limited QoS service and have lower priority compared with the UEs in the CSG.

## VII. PAGING PROCEDURE

Due to the high volume and small-sized femtocell deployment, it is well-known that paging messages is a big burden for the femtocell system. Figure 5 shows a paging procedure invoked by the UE with IMSI\_1. In the normal approach as with a macrocell network, the CN will only distinguish the paging area 1 where the paged IMSI is located. The femtocell-GW will send the paging message to all of the femtocells in paging area 1, which may cause a huge signaling redundancy for the large number of femtocells involved. One of the requirements agreed in [6] is that 'Additional registration and paging load as a result of HNB/HeNB deployment shall be minimized'. This means that paging optimization, namely minimizing the amount of paging messages used to page a UE in femtocells, is a confirmed requirement in 3GPP.

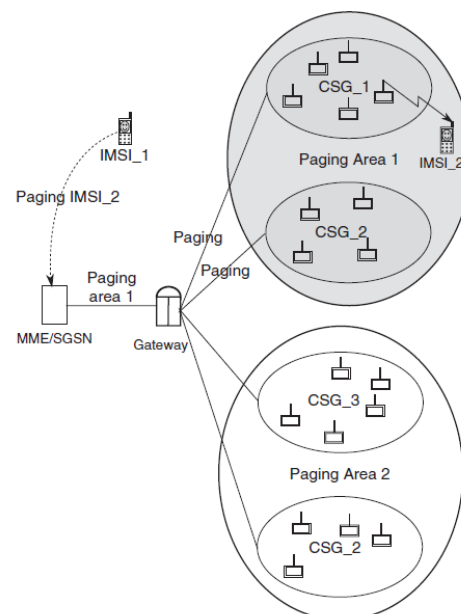


Fig. 5 Paging procedure with femtocells

### A. Cell Selection and Reselection

Cell selection and reselection are two basic mobility

procedures in wireless mobile networks. Cell selection takes charge of the selection of a suitable cell to camp on when it is powered on or after having previously lost coverage. Cell reselection however, enables UE to select a new serving cell when it meets the cell reselection criteria. Cell selection and reselection in the femtocell environment are more complicated than in a macrocell network.

### B. Cell Selection in Pre-release 8

Since cell selection is carried out irrespective of the neighbour cell list and will only select the cell with the best received signal as the serving cell, cell selection with femtocells is not a problem in pre-release 8 UTRAN. The policies used in macrocells can be applied in supporting cell selection with femtocells. However, a user may want to prioritize the femtocell over macrocell when the mobile is powered on or loses the coverage from the serving cell. Since Access Stratum (AS) will perform the PLMN search to select the best PLMN before it can pick up the strongest cell of that PLMN and Radio Access Technology (RAT) for the UE, different priorities for macrocell and femtocell can be achieved by assigning separated PLMN ID to them. In this case, the PLMN ID of the femtocell system will be assigned a higher priority than the macrocell layer. It should be noted that an additional PLMN ID is required while some operators may not be able to provide it. Two methods can be used here:

In Manual PLMN Selection, it enables the user to select the preferred mobile network from the available PLMN list. When separating the PLMN IDs of macrocell and femtocell, the user can manually select either the macrocell or femtocell network. Once the PLMN is selected the user will stick to that network and no mobility procedure will occur between these two PLMNs.

In National Roaming, the mobile is not roaming in its Home PLMN (HPLMN), but on a Visited PLMN (VPLMN) of the same country as the HPLMN. An HPLMN always has a higher priority than a VPLMN. In the automatic PLMN selection mode, the UE can be configured to search periodically for its HPLMN. The range of the HPLMN search timer is from 6 mins up to 8 hours with a default value of 60 mins [7].

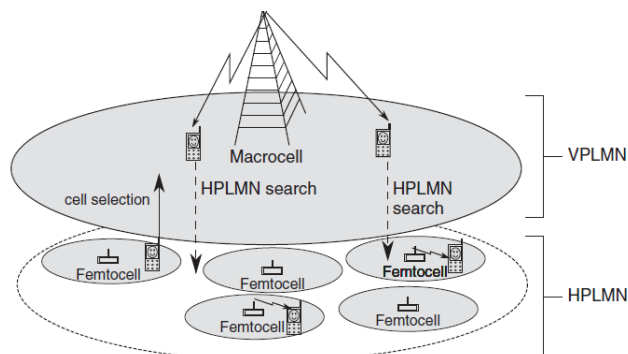


Fig. 6 Cell selection by national roaming (femtocell as HPLMN) in pre-release 8

One possible femtocell deployment scenario is to consider the femtocell layer as HPLMN and the macrocell layer as

VPLMN as shown in Figures 6 and 7. By setting a proper HPLMN search timer (i.e. 6 mins), the UE can camp on the femtocell as soon as it reaches the coverage of the femtocell and prioritizes it thereafter. However, the UE has to search for the femtocell PLMN periodically even when he is being served by a macrocell, which will drain the mobile battery. Besides, the operator may have set the UTRAN macrocell network as HPLMN, which may significantly impact the operators' service category when switching it to VPLMN.

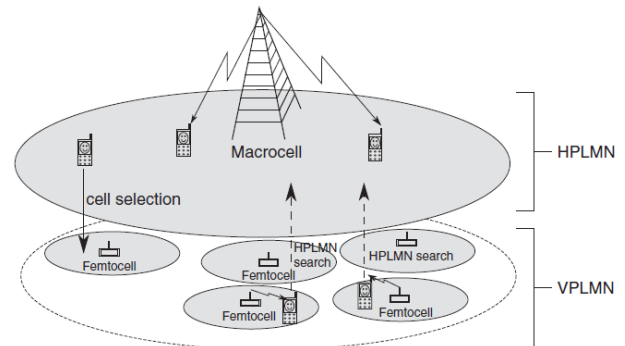


Fig. 7 Cell selection by national roaming (femtocell as VPLMN) in pre-release 8

The other femtocell deployment scenario is to consider the macrocell layer as HPLMN and the femtocell layer as VPLMN as shown in Figure 7. The UE will not search for the femtocell PLMN all the time when camped on the macrocell. However, as long as the UE is able to receive signals from the macrocell, it will be very difficult for the UE to select the femtocell even when the user receives a stronger signal from the femtocell.

In addition, when the UE is camping on the femtocell, it will periodically search for macrocells, which will also drain the battery.

### C. Cell Reselection in Pre-release 8

Cell reselection with femtocells is more complex than cell selection. During cell reselection, the UE needs to carry out cell measurements for intra-frequency, inter-frequency and inter-RAT neighbour cells and rank the cells based on a specified policy. Then the UE reselects the most suitable cell from the current serving cell. While dealing with femtocells in pre-release 8 UTRAN, to provide a reliable and clean neighbour cell list for a macrocell is still a problem as stated above. Thus, some operators/vendors do not implement the cell reselection features for the macrocell to femtocell reselection procedure (femtocell to macrocell or femtocell to femtocell reselection is easy, since the NCL of a femtocell will not be large and PSC confusion is unlikely to happen). Macrocell to femtocell mobility only relies on the cell selection procedure, which means users will have difficulty in selecting femtocell when under the coverage of macrocell. Nevertheless, there are still some operators/vendors who use their own method of NCL approach to support cell reselection from macrocell to femtocell.

Another important issue for cell reselection with femtocells is to prioritize allowed femtocells over macrocells, and avoid

unauthorized femtocells during the cell reselection procedure. Due to the lack of standardization of femtocell mobility procedures in pre-release 8 UTRAN, operators/vendors may realize cell reselection policies with their own substitution methods.

### VIII. CELL HANDOVER

Cell handover enables the UE to transfer the service seamlessly from its serving cell to the target cell without terminating the service. Since there is no Iur interface for femtocells in UTRA, soft handover is not supported by femtocells. Instead, hard handover is used. In order to support cell handover with femtocells, the issues with NCL, PSC/PCI confusion should be resolved as described in previous sections. Furthermore, the UE should avoid disrupting the service and try to keep the continuity of the service, making it even more complex than cell selection and reselection.

In pre-release 8 UTRA, a UE should initiate the cell measurement of all the neighbouring cells and the serving cell will make the handover decision based on the UE measurement report during cell handover procedure. In order to let the UE recognize the surrounding cells that need to be measured, the information on the neighbour cells needs to be included in the measurement control system information. For this purpose, the NCL is needed.

### IX. CONCLUSION

Femtocells are designed for very high quality wireless-connectivity inside home/office. So it's a challenge to control all the femtocell access points (FAPs) in an uncontrolled customer premises. From above all the methods and the type of processes, we can easily say that more research work is needed in this area to make the mobility management more efficient and easy between femtocell to macrocell, macrocell to femtocell, and femtocell to femtocell.

### REFERENCES

- [1] I. F. Akyildiz, J. McNair, J. S. M. Ho, H. S. Uzunalioglu, and W. Wang "Mobility Management in Next-Generation Wireless Systems" *proceedings of the IEEE*, vol. 87, pp. 8, August 1999.
- [2] 3GPP TR 25.820, "3G Home NodeB Study Item Technical Report" *3GPP-TSG RAN*, Valbonne, France, 8.2.0, Sep. 2008.
- [3] Motorola, "PCID confusion" *3GPP-TSG RAN*, Seoul, Korea, Tech. Rep. R2-092307, Mar. 2009.
- [4] 3GPP TS 25.304, "User Equipment (UE) procedures in idle mode and procedures for cell reselection in connected mode" *3GPP-TSG RAN*, 8.5.0, Mar. 2009.
- [5] 3GPP TS 36.304, "User Equipment (UE) procedures in idle mode" *3GPP-TSG RAN*, 8.5.0, Mar. 2009.
- [6] 3GPP TS 22.220, "Service requirements for Home NodeBs and Home eNodeB" *3GPP-TSG SA*, 9.0.0, Mar. 2009.
- [7] 3GPP TS 22.011, "Service accessibility" *3GPP-TSG SA*, bb, Tech. Rep. 9.1.0, Mar. 2009.
- [8] InterDigital, "Inbound handover to CSG and hybrid cell" *3GPP-TSG RAN*, Seoul, Korea, Tech. Rep. R2-092142, Mar. 2009.
- [9] V. Chandrasekhar, and J. G. Andrews, "Femtocell Networks: A Survey", *Communications Magazine, IEEE*, vol. 46, issue-9, pp. 59-67, September 2008.



# A Survey on Cryptographic based approach for Privacy Preserving Data Mining

A. Sweta R. Garasia

**Abstract**— The problem of privacy-preserving distributed data mining (PPDDM) overlaps closely with a field in cryptography for determining secure multi-party computations. The broad approach to cryptographic methods tends to compute functions over inputs provided by multiple recipients without actually sharing the inputs with one another. In this paper, I present a detailed overview and classification of approaches which have been applied for Secure Multiparty Computation (SMC).

**Index Terms**— Cryptography, Privacy Preserving Data Mining, SMC, Trusted Third Party

## I. INTRODUCTION

IN contrast to the centralized model, distributed Data Mining (DDM) model assumes that the data sources are distributed across multiple sites. Privacy-preserving data mining considers the problem of running data mining algorithms on confidential data that is not supposed to be revealed even to the party running the algorithm. The main aim of most distributed methods for privacy-preserving data mining is find useful information over the entire data set without compromising the privacy of the individual within the different participants.

In the following example, I show the necessity of privacy of individual within the distributed environment in real life application. In the field of medicine, consider the case that a number of different hospitals wish to clustering diseases, cures for diseases or symptoms of diseases. For that purpose they require patient data from different hospitals for the medical research. Furthermore, let us assume that privacy policy prevents these hospitals from revealing their data to each other, due to the confidentiality of patient records. In such case, different hospitals can not disclose their data directly. Rather it is necessary that they have to find solutions that enable the hospitals to compute the common function on the union of their database without revealing their privacy.

The rest of this paper is organized as follows: - In section 2, I discuss Theoretical background of Privacy Preserving Data Mining, while section 3 contains approaches of DDM.

In section 4, I have discussed SMC in detail. Section 5, conclude my study.

A. Asstt. Prof., Computer Engineering Department, Gandhinagar Institute of Technology, Moti Bhojan 382721, Gujarat, India (e-mail: sweta.garasia@git.org.in).

## II. THEORETICAL BACKGROUND

Here in this section first I describe the Privacy preserving data mining techniques. Then I will explore Distributed data mining in detail which contains theoretical background of secure multi party computation.

### A. Privacy preserving data mining

#### 1) Heuristic-based Approach

These techniques modify selected values i.e. changing some data values in a given dataset from an original value to another value [1].

#### 2) Reconstruction-based Approach

In these techniques the original distribution of the data is reconstructed. These algorithms are implemented by perturbing the data first and then reconstructing the distributions. It is important to realize that data modification results in degradation of the database performance [1].

#### 3) Cryptography-based Approach

Techniques like secure multiparty computation where a computation is secure if at the end of the computation, no party knows anything except its own input and the results [1]. SMC is mainly uses in distributed environment. Fig. 1 shows the taxonomy of Privacy Preserving Data Mining (PPDM) algorithms. My work is related to SMC.

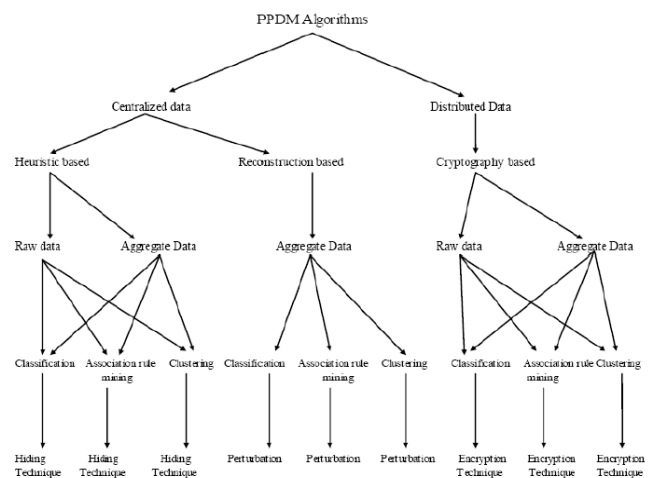


Fig 1. Taxonomy of the developed PPDM algorithms.

### B. Distributed Data Mining

Algorithms developed within this field focus are on efficiency, most of the algorithms developed to date do not take security consideration into account. A simple approach to data mining over multiple sources that will not share data is to run existing data mining tools at each site independently and combine the results [2]-[4] with distributed data, the way the data is distributed also plays an important role. In distributed data mining data are distributed in two ways as describe below. (i) Vertical partitioning [5], [6] of data implies that though different sites gather information about the same set of entities, they collect different feature sets. (ii) In horizontal partitioning [7] different sites collect the same set of information, but about different entities.

## III. APPROACHES OF DDM

The area of privacy preserving data mining has been extensively studied by the data mining community. In this discussion, I classify Distributed Data Mining algorithms into two categories.

### A. Data Perturbation based privacy

These algorithms aim at distorting the original private data, when released, do not reveal any individual's identifiable information. The first category of approaches introduces noise adding to the data or suppressing certain values, data transformation and reducing the granularity of the data to achieve partial privacy [8]. According to [9] perturbation techniques mix additive or multiplicative noise with the data so that actual values in the data set are not learned, yet the data mining results gathered from the perturbed data will not deviate significantly from the results gathered from the original data. Furthermore, algorithms in this category do not preserve privacy in any formal cryptological sense, i.e. one cannot easily calculate how much effort and resources are needed to filter out the noise and breach privacy. In additive perturbation [10], there is a private data set  $D = d_1, d_2, \dots, d_n$  and to every  $d_i \in D$  random noise  $r_i$  is added, where  $r_i$  is drawn from a known distribution such as a uniform distribution or a Gaussian distribution. The modified data set  $D' = d_1 + r_1, d_2 + r_2, \dots, d_n + r_n$  is released to the data miner. The data miner uses an expectation maximization algorithm to extract the values of  $d_i$  from  $d_i + r_i$ . The advantage of randomization based data Perturbation techniques is that privacy of the data can be preserved during the data collection process since the amount of noise to be added to each data record is independent of the later observations. In the perturbation approach, the distribution of each data dimension reconstructed independently. This means that any distribution based data mining algorithm works under an implicit assumption to treat each dimension independently. In many cases, a lot of relevant information for data mining algorithms such as classification is hidden in inter-attribute correlations. For example, the classification technique uses a distribution-based analogue of single-attribute split algorithm. However, other techniques such as multivariate decision tree algorithms cannot be accordingly modified to work with the perturbation

approach. This is because of the independent treatment of the different attributes by the perturbation approach.

### B. Cryptography based privacy

Cryptographic protocols are called private when their execution does not reveal any additional information about the involved parties' data, other than what is computed as a result of the protocol execution. This is done using the well known cryptographic protocol of Secure Multiparty Computation SMC [11]. SMC facilitates a group of people, each with its own private data, to perform some common computation task on the aggregate of their data. SMC ensures that, in the process, no personal information of data is revealed to anyone. However, the SMC based protocols are found to be extremely computationally expensive. The broad approach to cryptographic methods can be listed as either data encryption or secure multi-party computations.

## IV. SECURE MULTI-PARTY COMPUTATION

Privacy preserving distributed data mining requires multiple parties to collaborate for computing joint functions on their privately held data, while providing a privacy guarantee that the participants would not learn any information beyond what is implied by the output of the function computation. This is what even secures multi-party computation deals with. If there are  $n$  parties involved in a distributed data mining protocol where the  $i^{\text{th}}$  party owns data  $x_i$ , then secure multi-party computation is the approach to compute the function  $f$  on all parties' data  $f(x_1, x_2, \dots, x_n) = (y_1, y_2, \dots, y_n)$ , such that party  $i$  only gets to know  $y_i$  and nothing else. An example of such a computation is Yao's millionaire problem [12]. The problem description is as follows: two millionaires meet in the street and want to find out who is wealthier without having to reveal their actual fortune to each other. The function computed in this case is a simple comparison between two numbers. If the result is that the first millionaire is wealthier, then he knows that, but this should be all the information he learns about the other guy and not the exact value of his assets.

### A. Models in SMC

Now I cover some of the different models of computation in SMC [6].

#### 1) Trusted Third Party Model

The gold standard for security is the assumption that we have a trusted third party to whom we can give all data. The third party performs the computation and delivers only the results except for the third party, it is clear that nobody learns anything not inferable from its own input and the results. The goal of secure protocols is to reach this same level of privacy preservation, without the problem of finding a third party that everyone trusts.

#### 2) Semi-honest Model

The Semi-honest model is also known in the literature as the honest-but-curious model. A semi-honest party follows the rules of the protocol using its correct input, but after the protocol is free to use whatever it sees during execution of the protocol to compromise security / privacy.

*Composition Theorem* for the semi-honest model. Suppose that  $g$  is privately reducible to  $f$  and that there exists a protocol for privately computing  $f$ . Then there exists a protocol for privately computing  $g$ . Informally, the theorem states that if a protocol is shown to be secure except for several invocations of sub-protocols, and if the sub-protocols themselves are proven to be secure, then the entire protocol is secure. The immediate consequence is that, with care, we can combine secure sub-protocols to produce new secure protocols. Also, if many algorithms depend on a few common sub-protocols, efficient implementation of these sub-protocols significantly improves the overall efficiency.

### 3) Malicious Model

In the malicious model, no restrictions are placed on any of the participants. Thus any party is completely free to spoil in whatever actions it please.

## B. Protocols of PPDDM

Here in this section I will explore the sub protocols used in SMC.

### 1) Oblivious Transfer Protocol

A key building-block for many kinds of secure function evaluations is the 1 out of 2 oblivious-transfer protocol [13][14] which involves two parties: a sender, and a receiver. The sender's input is a pair  $(x_0, x_1)$  and the receiver's input is a bit  $b \in \{0, 1\}$  denoting the index of  $x_b$ . At the end of the protocol the receiver learns  $x_b$  and nothing else, and the sender learns nothing. There can be many ways for implementing the oblivious transfer protocol. One simple way is for the receiver to generate two random public keys,  $K_0$  and  $K_1$ , but to know only the decryption key for  $K_b$ . Using the public keys the sender can encrypt  $(x_0, x_1)$  and send it back to the receiver who can decrypt only one of them  $x_b$  using the decryption key. Oblivious transfer is sufficient for secure computation in the sense that, given an implementation of oblivious transfer, it is possible to securely evaluate any polynomial-time computable function without any additional primitive. Oblivious transfer can be used to design secure protocols for both semi-honest and malicious adversaries and there exist generalizations of the 1 out of 2 oblivious protocols to 1 out of  $N$  oblivious protocol in the literature for designing efficient secure function computations.

### 2) Circuit Evaluation

Yao [12] presents a constant round protocol for secure computation of probabilistic polynomial time functions by expressing the functions as combinatorial circuits with gates defined over some fixed base. The polynomial size circuit consists of AND and XOR gates and the input bits are transmitted through wires connecting these gates. The protocol requires one of the parties to generate an encrypted or "garbled" circuit representing the function to be evaluated,  $f$ , and send it to the other party. The receiver can then reconstruct the values from the garbled representation using a 1 out of 2 oblivious transfer protocol. Using this information the receiver can now compute the output of the circuit himself. Although Yao's generic circuit evaluation method is secure, it poses significant computational problems since the computational complexity of the protocol is roughly linear in

relation to the size of the input and the communication complexity is linear in relation to the size of the circuit. Given the size and computational cost of data mining problems, representing algorithms as a Boolean circuit results in unrealistically large circuits. Therefore, this technique is not used usually for distributed privacy preserving data mining problems. A number of secure multiparty computation protocols have been adopted for different privacy preserving data mining tasks till date. A classic problem which is often used as a primitive for many other problems in data mining is that of computing the scalar dot product in a distributed environment and Du and Atallah [15] describe a systematic set of methods for transforming a number of privacy preserving data mining problems into secure inner product computation.

## V. CONCLUSION

Secure multi-party computation is very relevant to the line of research involving privacy reserving data mining in distributed. This paper was intended to demonstrate basic ideas from a large body of cryptographic research on secure distributed computation.

## REFERENCES

- [1] T. Mielikainen. "On inverse frequent set mining". In Proceedings of the 3rd IEEE ICDM Workshop on Privacy Preserving Data Mining, *IEEE Computer Society*, pp.18–23, 2003.
- [2] P. Chan. "An Extensible Meta-Learning Approach for Scalable and Accurate Inductive Learning". PhD thesis, Department of Computer Science, Columbia University, New York, NY, 1996
- [3] P. Chan. "On the accuracy of meta-learning for scalable data mining". *Journal of Intelligent Information Systems*, 8:5{28, 1997}
- [4] A. Prodromidis, P. Chan, and S. Stolfo. "Advances in Distributed and Parallel Knowledge Discovery", *AAAI/MIT Press*, September 2000.
- [5] J. Vaidya. Thesis on Privacy Preserving data mining over vertical partitioned data, the Faculty of Purdue university, August 2004
- [6] C. Aggarwal. "Privacy-Preserving Data Mining Models and Algorithms" *IBM T.J. Watson Research Center*, US and Philip S. Yu University of Illinois at Chicago, US 2008.
- [7] D. Cheung, V. Ng, A. Wai-Chee Fu, and Y. Fu. "Efficient mining of association rules in distributed databases". *IEEE Transactions on Knowledge and Data Engineering*, 8(6):911{922} December 1996.
- [8] K. Kargupta, H. Datta, S. Wang, Q. Sivakumar, "On the privacy preserving properties of random data perturbation techniques". In: *ICDM*, pp. 99–106,2003
- [9] M. Doganay , T. Pedersen , Y. Saygin , E. Savaş , A. Levi. "Distributed privacy preserving k-means clustering with additive secret sharing", Proceedings of the 2008 international workshop on *Privacy and anonymity in information society*, March 29-29, 2008
- [10] R. Agrawal and R. Srikant, "Privacy-preserving data mining". In Proceedings of the *2000 ACM SIGMOD International Conference on Management of Data*, May 16-18, 2000, Dallas, Texas, USA, pages 439–450, ACM 2000
- [11] O. Goldreich, "The Foundations of Cryptography", vol. 2. *Cambridge Univ. Press*, Cambridge,2004
- [12] Yao, Andrew C. "How to generate and exchange secrets". In *Proceedings of the 27th IEEE Symposium on Foundations of Computer Science*, pages 162–167.1986
- [13] Yao, Andrew C. "How to generate and exchange secrets". In *Proceedings of the 27th IEEE Symposium on Foundations of Computer Science*, pages 162–167.1986
- [14] S. Even, O. Goldreich, and A. Lempel, "A randomized protocol for signing contracts". *Communications of the ACM*, 28(6):637–647, 1985.
- [15] W. Du and M. Atallah. "Secure multi-party computation problems and their applications: A review and open problems". In Proceedings of the

2001 Workshop on *New Security Paradigms*, Cloudcroft, NM, pages  
13–22, ACM September 2001

# A Survey Paper on Hyperlink-Induced Topic Search (HITS) using networks intelligence

A. Mr.Ramesh Prajapati

**Abstract--**In this Paper Hyperlink-Induced Topic Search (HITS) is a link analysis algorithm using networks intelligence which helps in rating Web pages also known as Hubs and authorities fast and efficient HITS mechanism for web crawling and retrieval remains as a challenging issue. This paper deals with analysis and comparison of web page ranking algorithms based on various parameter to find out their advantages and limitations for the ranking of the web pages using Network intelligence. The primary goal of the web site owner is to provide the relevant information to the users to fulfill their needs. Web mining technique is used to categorize users and pages by analyzing users behavior, the content of pages and order of URLs accessed. In this paper we discuss and compare the different used algorithms i.e. Page Rank and HITS.

**Index Terms—**Hub, HITS, Networking, Page Rank, Web Mining, Web Usage Mining, Web Content Mining, Web Structure Mining, Web Usage Mining, Weighted Page Rank, HITS

## I. INTRODUCTION

As the volume of information on the internet is increasing Day by day so there is a challenge for website owner to Provide proper and relevant information to the internet user. Retrieving of the required web page on the web, efficiently and effectively, is becoming a challenge. Whenever a user wants to search the relevant pages, he/she prefers those relevant pages to be at hand. The bulk amount of information becomes very difficult for the users to find, extract, filter or evaluate the relevant information. This issue raises the necessity of some technique that can solve these challenges. Web mining can be easily executed with the help of other areas like Database (DB), Information retrieval (IR), Natural Language Processing (NLP), and Machine Learning etc. These can be used to discuss and analyze the useful information from WWW.

Following are some challenges:

1) Web is huge. 2) Web pages are semi structured. 3) Web information stands to be diversity in meaning. 4) Degree of quality of the information extracted. 5) Conclusion of knowledge from information extracted

A .Lecturer, Department of Information Technology, Gandhinagar Institute of Technology, [ramesh.prajapati@git.org.in](mailto:ramesh.prajapati@git.org.in)

This paper is organized as follows- Web Mining is introduced in Section II. The areas of Web mining i.e. Web Content Mining, Web Structure Mining and Web Usage Mining are discussed in Section III. Section IV describe the Scale-free network model. In Section V describes the various Links based Analysis algorithms. Page Rank algorithm and its Limitation of PageRank are presented in Section A. In Section B includes Weighted PageRank algorithms. In Section C HITS, Hub and Authorities and Motivation behind HITS. HITS Algorithm and Handling “spam” links In Section D. Handling Span links in Section E. Section VI Based on the literature analysis provides the comparison of HITS vs. PageRank algorithms. Concluding remarks are given in Section VII.

## II. WEB MINING

Web mining is the Data Mining technique that automatically discovers or extracts the information from web documents. It is the extraction of interesting and potentially useful patterns and implicit information from activity related to the World.

### A. Web Mining Process

The complete process of extracting knowledge from Web data [2] is follows in Fig.1:

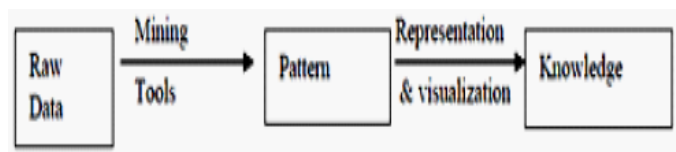


Figure 1: Web Mining Process

It consists of following tasks [4]:

1. Resource finding: It involves the task of retrieving Intended web documents.
2. Information selection and pre-processing: It involves the automatic selection and pre processing of specific information from retrieved web Resources.
3. Generalization: It automatically discovers general Patterns at individual web sites as well as across multiple sites.

4. Analysis: It involves the validation and interpretation of the mined patterns. A human plays an important role in information on knowledge discovery process on web.

### III. WEB MINING CATEGORIES

There are three areas of Web Mining according to the web data used as input in Web Data Mining. Web Content Mining, Web Structure Mining and Web Usage Mining.

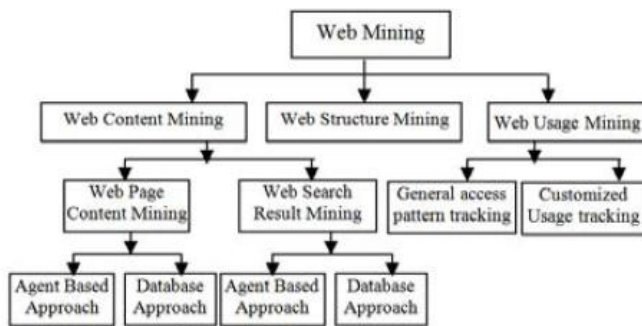


Figure 2: Classification of Web Mining

#### A. Web Content Mining

It is the process of retrieving the information from Web document into more structured forms and indexing the information to retrieve it quickly. It focuses mainly on the structure within a document i.e. inner document level. Web Content Mining is related to Data Mining because many Data Mining techniques can be applied in Web Content Mining. It is also related with text mining because much of the web contents are text, but is also quite different from these because web data is mainly semi structured in nature and text mining focuses on unstructured text.

#### B. Web Structure Mining

It is the process by which we discover the model of link structure of the web pages. We catalog the links; generate the information such as the similarity and relations among them by taking the advantage of hyperlink topology. The goal of Web Structure Mining is to generate structured summary about the website and web page. Page Rank and hyperlink analysis also fall in this category. It tries to discover the link structure of hyper links at inter document level. As it is very common that the web documents contain links and they use both the real or primary data on the web so it can be concluded that Web Structure Mining has a relation with Web Content Mining. It is using the tree-like structure to analyze and describe the HTML (Hyper Text Markup Language).

#### C. Web Usage Mining

It is the process by which we identify the browsing patterns by analyzing the navigational behavior of user. It focuses on techniques that can be used to predict the user behavior while the user interacts with the web. It uses the secondary data on the web. This activity involves the

automatic discovery of user access patterns from one or more web servers. Through this mining technique we can ascertain what users are looking for on Internet. It consists of three phases, namely preprocessing, pattern discovery, and pattern analysis. Web servers, proxies, and client applications can quite easily capture data about Web usage.

### IV. SCALE-FREE NETWORK MODEL

A simple model for generating “scale-free” networks in following point.

1. Evolution: networks expand continuously by the addition of new vertices, and
2. Preferential-attachment (rich get richer): new vertices attach preferentially to sites that are already well connected. Growing the network (evolution): Starting with a small number ( $m_0$ ) of vertices, at every time step we add a new vertex with  $m$  ( $\leq m_0$ ) edges that link the new vertex to  $m$  different vertices already present in the system. Growing the network (preferential attachment): To incorporate Preferential attachment, we assume that the probability  $P$  that a new vertex will be connected to vertex  $i$  depends on the Connectivity  $k_i$  of that vertex, so that  $P(k_i) = k_i / \sum k_j$ .

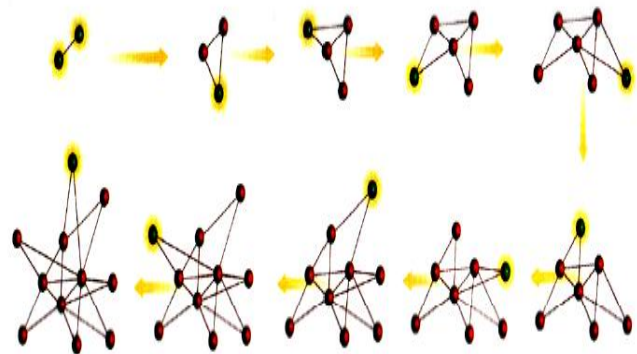


Figure 3: Scale-free network model

### V. LINK BASED ANALYSIS

Web mining technique provides the additional information through hyperlinks where different documents are connected. We can view the web as a directed labeled graph whose nodes are the documents or pages and edges are the hyperlinks between them. This directed graph structure is known as web graph. There are number of algorithms proposed based on link analysis. Three important algorithms Page Rank, Weighted PageRank and HITS are discussed below.

#### A. Page Rank Algorithm

Page Rank is a numeric value that represents how important a page is on the web. Page Rank is the Google’s

method of measuring a page's "importance." When all other factors such as Title tag and keywords are taken into account, Google uses Page Rank to adjust results so that more "important" pages move up in the results page of a user's search result display. Google Fig.s that when a page links to another page, it is effectively casting a vote for the other page. Google calculates a page's importance from the votes cast. For it. Its provides a better approach that can compute the importance of web page by simply counting the number of pages that are linking to it. These links are called as backlinks.If a backlink comes from an important page than this link is given higher weightage than those which are coming from non-important pages. The link from one page to another is considered as a vote. Not only the number of votes that a page receives is important but the importance of pages that casts The algorithm of Page Rank as follows:

Page Rank takes the back links into account and propagates the ranking through links. A page has a higher rank, if the sum of the ranks of its backlinks is high. Fig. 3 shows an example of back links wherein page A is a backlink of page B and page C while page B and page C are backlinks of page D.The original Page Rank algorithm is given in following equation

$$PR(P)=(1-d)+d(PR(T1)/C(T1)+.....PR(Tn)/C(Tn)) \dots (1)$$

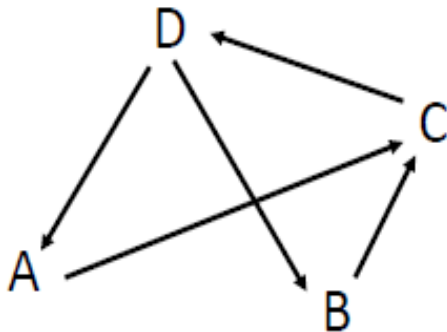


Figure 4: Backlinks Page Rank

Where, PR (P)= PageRank of page P  
 PR (Ti) = PageRank of page Ti which link to page  
 C (Ti) =Number of outbound links on page T  
 D = Damping factor which can be set between 0 and 1.

A link from A to B is a vote for B cast by A. Votes cast by pages that are *important* weigh more heavily. But there are different types of important nodes:

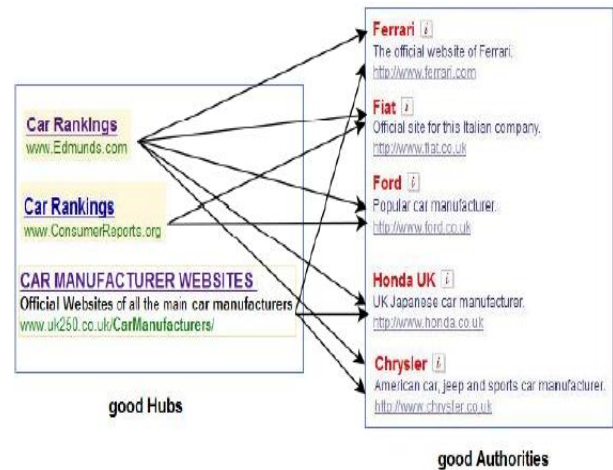


Figure 5: Top automobile markers

Problem to be solved relevant terms may not appear on the pages of authoritative websites. Many prominent pages are not self descriptive. Car manufacturers may not use the term “automobile manufacturers” on their home page. The term “search engine” is not used by any of natural authorities like Yahoo, Google, and AltaVista etc.

*B. Weighted Page Rank*

Extended Page Rank algorithm- Weighted Page Rank assigns large rank value to more important pages instead of dividing the rank value of a page evenly among its outlink pages. The importance is assigned in terms of weight values to incoming and outgoing links denoted as and respectively. This algorithm was proposed by Wenpu Xing and Ali Ghorbani which is an extension of PageRank algorithm. This Algorithm assigns rank values to pages according to their importance rather than dividing it evenly. The importance is assigned in terms of weight values to incoming and outgoing links.

This is denoted as  $W^{in}(m, n)$  and  $W^{out}(m,n)$  respectively.  $W^{in}(m, n)$  is the weight of link(m,n) as given in . It is calculated on the basis of number of incoming links to page n and the number of incoming links to all reference pages of page m.

$$W^{in}_{(m,n)} = \frac{I_n}{\sum_{p \in R(m)} I_p} \dots(2)$$

$I_n$  is number of incoming links of page n,  $I_p$  is number of incoming links of page p,  $R(m)$  is the reference page list of page m.  $W^{out}(m,n)$  is the weight of link(m,n)as given in (3). It is calculated on the basis of the number of outgoing links of page n and the number of outgoing links of all the reference pages of page m.

$$W_{(m,n)}^{out} = \frac{O_n}{\sum_{p \in R(m)} O_p} \dots(3)$$

$O_n$  is number of outgoing links of page n,  $O_p$  is number of outgoing links of page p, Then the weighted PageRank is given by formula in (4)

$$WPR(n) = (1 - d) + d \sum_{m \in B(n)} WPR(m) W_{(m,n)}^in W_{(m,n)}^{out} \dots(4)$$

C. HITS (Hyper-link Induced Topic Search)

Hyperlink-Induced Topic Search (HITS) (also known as Hubs and authorities) is a link analysis algorithm that rates Web pages, developed by Jon Kleinberg. It was a precursor to PageRank. The idea behind Hubs and Authorities stemmed from a particular insight into the creation of web pages when the Internet was originally forming; that is, certain web pages, known as hubs, served as large directories that were not actually authoritative in the information that it held, but were used as compilations of a broad catalog of information that led users directly to other authoritative pages. In other words, a good hub represented a page that pointed to many other pages, and a good authority represented a page that was linked by many different hubs. The scheme therefore assigns two scores for each page: its authority, which estimates the value of the content of the page, and its hub value, which estimates the value of its links to other pages. A page may be a good hub and a good authority at the same time. The HITS algorithm treats WWW as directed graph  $G(V,E)$ , where  $V$  is a set of vertices representing pages and  $E$  is set of edges corresponds to link. Attempts to computationally determine hubs and authorities on a particular topic through analysis of a relevant sub graph of the web.

Based on mutually recursive facts: Hubs point to lots of authorities. Authorities are pointed to by lots of hubs. **Authority:** A valuable and informative webpage usually pointed to by a large number of hyperlinks • **Hub:** A webpage that points to many authority pages is itself a resource and is called a hub • Authorities and hubs **reinforce** one another. A good authority is pointed to by many good hubs. A good hub points to many good authorities

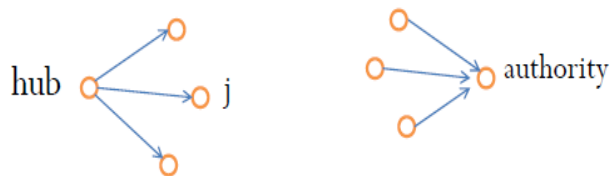


Figure 6: Hub and Authority

The creator of page p, by including a link to page q, has in some measure conferred authority on q Links afford us the opportunity to find potential authorities purely through the pages that point to them

What is the problem here? Some links are just navigational “Click here to return to the main menu” Some links are advertisements Difficulty in finding balance between relevance and popularity Solution: Based on relationship between the authorities for a topic and those pages that link to many related authorities-hubs.

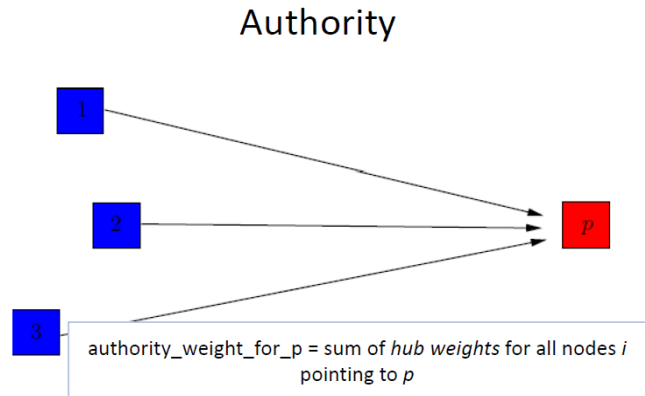


Figure 7: Relationship between Authority and Hub

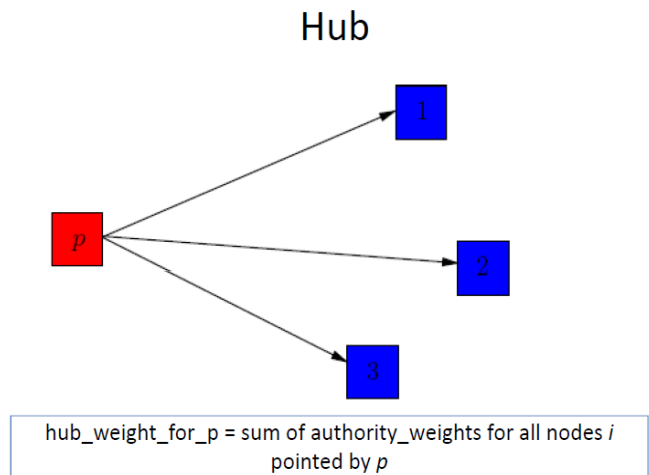


Figure 8: Relationship between Hub and Authority

D.HITS Algorithm

Computes hubs and authorities for a particular topic specified by a normal query. • First determines a set of relevant pages for the query called the *base set S*. Analyze the link structure of the web subgraph defined by *S* to find authority and hub pages in this set. Following point’s construction of focused sub graph.

Motivation behind HITS:

- We have a set created by text-based search engine.
- Why do we need subset?



- The set may contain too many pages and entail a Considerable computational cost
- Most of the best authorities may not belong to this set
- Subset properties:
- Relatively small
- Rich in relevant pages
- Contains most (or many) of the strongest authorities

## You need relevance – Start filtering

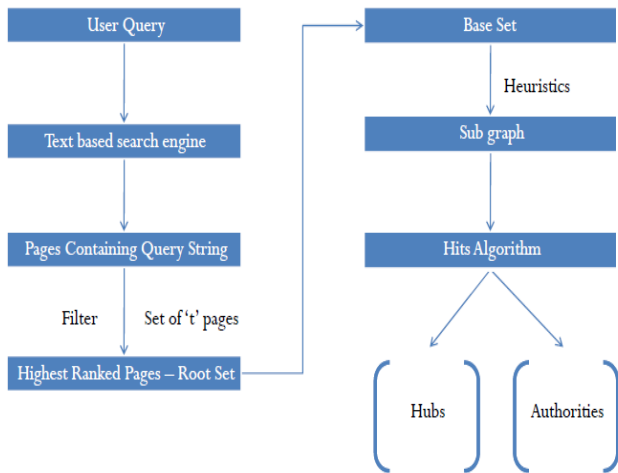


Figure 9: Computes hubs and authorities for a particular topic specified by a normal query

First find a set of relevant pages

- For a specific query  $Q$ , let the set of documents returned by a standard search engine be called the *root* set  $R$ .
- Initialize  $S$  to  $R$ .
- Add to  $S$  all pages pointed to by any page in  $R$ .
- Add to  $S$  all pages that point to any page in  $R$ .

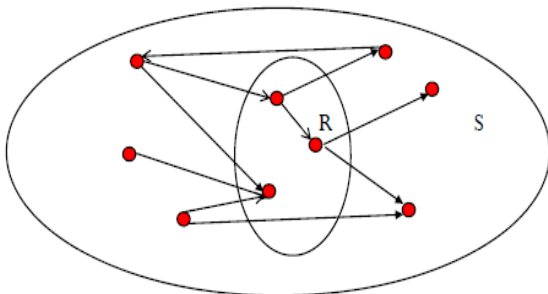


Figure 10: Determines a set of relevant pages for the query

The Subgraph reduction Offset the effect of links that serve purely a navigational function Remove all intrinsic edges from the graph, keeping only the edges corresponding to transverse link. Remove links that are mentioned in more than  $m$  pages ( $m=4-8$ ).

Calculating the Hub and Authority Weights

- $A$  is the adjacency matrix of graph  $G=(V,E)$

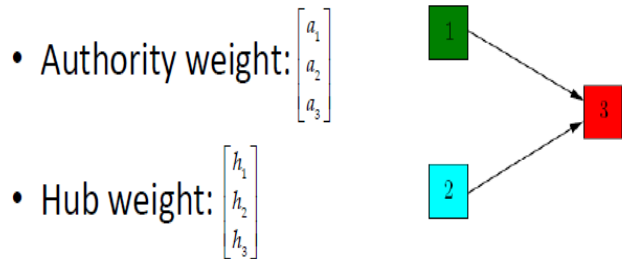


Figure 11: Calculating the Hub and Authority Weights

Iterative Algorithm

Each page  $p$  is assigned two non-negative weights, an authority weight  $a$  and a hub weight  $h$ . Update the weights of  $a$  and  $h$

– Authority weight: 
$$a(j) = \sum_{i:(i,j) \in E} h(i)$$

– Hub weight: 
$$h(j) = \sum_{i:(j,i) \in E} a(i)$$

These operations add the weights of hubs into the authority weight and add the authority weights into the hub weight respectively. Alternating these two operations will eventually result in an equilibrium value, or weight, for each page.

$G$ : a collection of  $n$  linked pages

- Set  $a_0 = [1/n, \dots, 1/n]^T$
- Set  $h_0 = [1/n, \dots, 1/n]^T$
- For  $t=1, 2, \dots, k$

– For  $j= 1, 2, \dots, n$

- Obtain new authority weights  $a'_t(j) = \sum_{i:(i,j) \in E} h_{t-1}(i)$
- Normalize weights  $a_t(j) = \frac{a'_t(j)}{\sum_j a'_t(j)}$
- Obtain new hub weights  $h'_t(j) = \sum_{i:(j,i) \in E} a_t(i)$
- Normalize weights  $h_t(j) = \frac{h'_t(j)}{\sum_j h'_t(j)}$

– end

- End

Based on the review of this HITS algorithm the overall graph of with Authority and Hubness represented as following. HITS algorithm discovered, they share similar roles in terms of their email communication pattern in the data set. Our algorithm discovers this structure as well. The estimated rankings are so close to the actual ones that it is difficult to distinguish them.



Figure 12:-Authority and Hubness Weights

**E. Handling “spam” links**

In this algorithm all links should be equally treated so we considerations two problem. Some links may be more meaningful/important than other links. Web site creators may trick the system to make their pages more authoritative by adding dummy pages pointing to their cover pages (spamming).then we implements following things.

Transverse link: links between pages with different domain names.

- Domain name: the first level of the URL of a page.
- Intrinsic link: links between pages with the same domain name.

Transverse links are more important than intrinsic links.

Two ways to incorporate this:

1. Use only transverse links and discard intrinsic links.
2. Give lower weights to intrinsic links.

How to give lower weights to intrinsic links? In adjacency matrix A, entry (i, j) should be assigned as follows:

- If i has a transverse link to j, the entry is 1.
- If i has an intrinsic link to j, the entry is c,

Where  $0 < c < 1$ .

- If i have no link to j, the entry is 0.

**VI. COMPARISON**

Table1 shows the difference between above two algorithms:

Table 1: Comparison of Page Rank and HITS

Algorithm	PageRank	HITS
Calculation	Calculates score after indexing process.	Calculate score without indexing.
Operates	Operates on a big web Graph focusing on all the back links and relevance factors.	on small sub graphs representing a linkage between Hub and Authority websites
I/P Parameters	Backlinks	Backlinks, forward links and content
Search Engine	Google	Clever
Use	in websites relational analysis specifically	In multiple environments from institutes to search engine crawlers.
Mining Technique Used	WSM	WSM and WCM
Working Procedure	focus on both authoritative pages and good hub pages	Focus on the authoritative pages.

**VII. CONCLUSION**

Web Mining is powerful technique used to extract the Information from past behavior of users. Various Algorithms are used in Web Structure Mining to rank the relevant pages. PageRank, Weighted PageRank and HITS treat all links equally when distributing the rank score. In the Problem of

page rank and weight page algorithm relevant terms may not appear on the pages of authoritative websites. Many prominent pages are not self descriptive. In HITS algorithm all links should be equally treated so we considerations two problem. Some links may be more meaningful/important than other links. Web site creators may trick the system to make their pages more authoritative by adding dummy pages pointing to their cover pages (spamming) network intelligence. We observe that the HITS algorithm effectively identifies the congestion hot spots in operational networks intelligence. It can identify the congestion hot spots with the help of the network topology alone. For the future works, there are still many issues that need to be explored. By using HITS method, for any network, the congestion hot spots can be calculated with minimum effort and requirements. With the HITS algorithm, we improve the quality of routing to almost perfection thereby avoiding congestion hot spots in the network.

Being HITS algorithms are not good enough to be applied in mining the informative structures, we brings forward a new CGHITS Web search algorithm based on hyperlinks and content relevance strategy, and the cause of topic drifting away in HITS algorithm is analyzed. CGHITS algorithm evolves a population of web pages for maximizing the relevance by clone, crossover and selection operator. Experiments are carried out to show that CGHITS algorithm was able to achieve the optimal solution in most cases

#### REFERENCES

- [1] Rekha Jain, Dr G.N.Purohit, "Page Ranking Algorithms for Web Mining," International Journal of Computer application, Vol 13, Jan 2011.
- [2] Cooley, R, Mobasher, B., Srivastava, J."Web Mining: Information and pattern discovery on the World Wide Web". In proceedings of the 9th IEEE International Conference on tools with Artificial Intelligence (ICTAI' 97).Newposrt Beach,CA 1997.
- [3] Pooja Sharma, Pawan Bhadana, "Weighted Page Content Rank For Ordering Web Search Result", International Journal of Engineering Science and Technology, Vol 2, 2010.
- [4] R. Kosala, H. Blockeel "Web mining research" A survey. ACM Sigkdd Explorations,2(1):1-15, 2000.
- [5] Wang jicheng, Huang Yuan,Wu Gangshan, Zhang Fuyan, "Web mining: Knowledge discovery on the Web Systems", Man and Cybernetics 1999 IEEE SMC 99 conference Proceedings. 1999 IEEE International conference
- [6] Raymond Kosala, Hendrik Blockee, "Web Mining Research : A Survey", ACM Sigkdd Explorations Newsletter, June 2000, Volume 2.
- [7] Taher H. Haveliwala, "Topic-Sensitive Page Rank: A Context-Sensitive Ranking Algorithms for Web Search", IEEE transactions on Knowledge and Data Engineering Vol.15, No 4 July/August 2003.
- [8] J. Hou and Y. Zhang, "Effectively Finding Relevant Web Pages from Linkage Information, IEEE Transactions on Knowledge and Data Engineering", Vol. 15, No. 4, 2003.
- [9] P Ravi Kumar, and Singh Ashutosh kumar," Web Structure Mining Exploring Hyperlinks and Algorithms for Information Retrieval, American Journal of applied sciences", 7 (6) 840-845 2010.
- [10] M.G. da Gomes Jr. and Z. Gong, "Web Structure Mining: An Introduction, Proceedings of the IEEE International Conference on Information Acquisition", 2005.
- [11] R. Kosala, and H. Blockeel," Web Mining Research: A Survey, SIGKDD Explorations, Newsletter of the ACM Special Interest Group on Knowledge Discovery and Data Mining" Vol. 2, No. 1 pp 1-15, 2000.

# FPGA Based Implementation of Huffman Coder using VHDL

A. Rutvij C.Joshi, B. Anil J. Kshatriya, c. V.K.Thakar

**Abstract-**This is one step forward towards the implementation of FPGA/CPLD based design using HDL(Hardware description language). This paper introduces an idea to implement a Static Huffman coder on FPGA using Hardware description language. For the, secure and reliable data communication, efficient coding-decoding mechanism is necessary. Traditionally TTL devices are used to implement an embedded coding-decoding mechanism but now FPGA based designs are become more popular because of its certain very important advantages like High speed, low power consumption, and Small size. The VHDL has been used to develop logic and that logic is implemented using FPGA. The sole purpose of this embedded system is to encode the source data according to its probability of occurrence

in to the communication channel. The entire system is implemented on FPGA prototyped board using VHDL.

## I. INTRODUCTION

Source coding was introduced long back by Shannon , and there are many source coding techniques available, but among them Huffman[1] coding technique is one of the best coding technique. This paper introduces CMOS devices like FPGAs/CPLDs to implement embedded source encoder using Hardware description language.

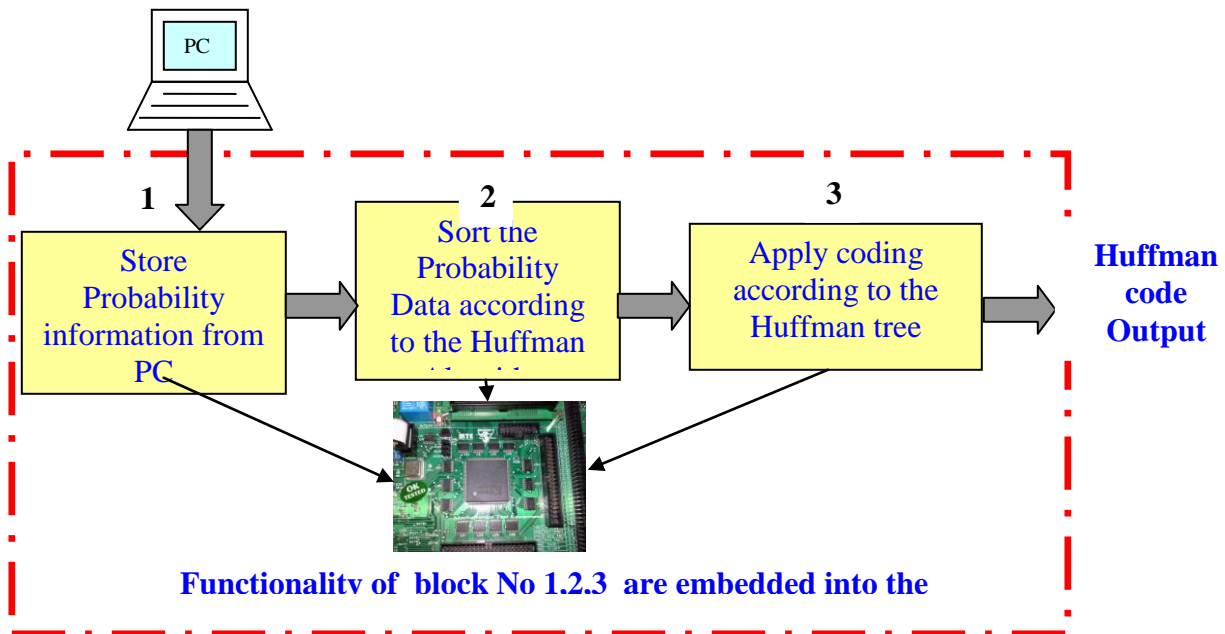


Fig.1 Block Diagram of FPGA based Huffman coder

The source coding theorem is one of the three fundamental theorems of information theory. The source coding theorem establishes a fundamental limit on the rate at which the output of an information source can be compressed without causing a larger error probability[2]. Source coding may be classified

into three areas Data Compression: to compress the data, to encrypt the information for Security purpose, to change the form of representation of data that it can be transmitted over communication Channel. In Information theory Huffman<sup>1</sup> coding is an entropy encoding algorithm used for lossless

data compression. In Huffman coding, fixed length blocks of the source outputs are mapped to variable length binary blocks. This is called fixed to variable length coding. The idea is to map the more frequently occurring fixed length sequences to shorter binary sequences and the less frequently occurring ones to longer binary sequences, so that overall bandwidth of the communication system can be optimized[10].

II. FPGA (FIELD PROGRAMMABLE GATE ARRAY)

During last decade there has been rapid development of many different types of reprogrammable hardware devices, such devices go by many names like reconfigurable hardware, reprogrammable hardware and so on[4]. As explain earlier traditionally TTL devices like Microprocessors, and Micro controllers are utilized for the development of automation systems but now days CMOS devices like FPGA and CPLD have prove their competency in the field of automation. FPGA is the acronym for Field Programmable Gate Array. FPGA is nothing but a bank of CLBs(Configured logic blocks) as shown into the Fig.2

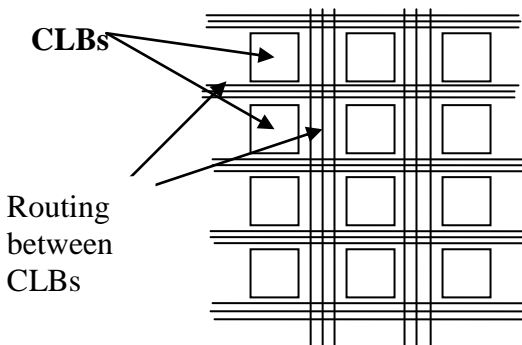


Fig.2 An array of CLBs<sup>3</sup>

The basic element of each CLB is CMOS. Architecture of CLB is vary from manufacture to manufacture. Fully fabricated FPGA chips containing tens to hundreds of thousand or even more logic gates with programmable interconnects, are available to users for their custom hardware programming to realize desired functionality. This design style provides a means of fast prototyping and also for cost-effective chip design. The interconnection between all configure logical blocks (CLB )can be accomplished by applying suitable program into the RAM cells. FPGA is software configured, so users can modify their design very easily, this results in significant cost saving in design and production[5]. In this work VHDL is used as a programming language and Project Navigator is used as synthesized tool. Computer is connected with FPGA using JTAG connector. FPGA can be reprogrammed easily using one of the VHDL

programming tools through computer. The Photograph of Mechatronics make Xilinx FPGA prototype board is shown in Fig.3



Fig.3 Photograph of Mechatronics make FPGA prototype board[12]

Important Features of FPGA[5]:

- [1] Instantaneous Implementation: As memory-writing times are usually short, designs can effectively be implemented "instantaneously".
- [2] Dynamic reconfiguration: Parts of the FPGA can be reprogrammed at run-time, that is, in the application situation.
- [3] Systems built using FPGAs can be upgraded in the field in order to correct bugs, to recover from damage, or to add new functionality.
- [4] Systems built using FPGAs can be upgraded in the field in order to correct bugs, to recover from damage, or to add new functionality.

III. HUFFMAN CODING

Huffman coding is one of the basic and most frequently employed lossless compression methods. Huffman coding uses a specific method for choosing the representation for each symbol, resulting in a prefix-free codes that expresses the most common character using shorter strings of bits than are used for less common source symbols. For a set of symbols with uniform probability distribution and a number of members which is a power of two, Huffman coding is equivalent to simple binary block encoding but for text data which contain lots of alphabets and numerical characters then probability distribution is not even for each alphabet. The sum of probability across all symbols is always less then or equal to one, if the summation of all probabilities are one than this type of code is known as a "complete code"[10]. In Huffman coding input of the system are set of symbols and their probability, and out come is the Huffman encoded data for each symbol.

For example suppose A is the set of Alphabets

and  $W$  is the set of Probabilities then Generated Huffman code is  $C$  for  $n$  –number of Symbols.

$$A = \{a_1, a_2, a_3, \dots, a_n\}$$

$$W = \{w_1, w_2, w_3, \dots, w_n\}$$

$$C(A, W) = \{c_1, c_2, c_3, \dots, c_n\}$$

The information content  $h$  of each symbol  $a_i$  with non zero probability

$$h(a_i) = \log_2 \frac{1}{w_i}$$

Average length of the Huffman code is define by

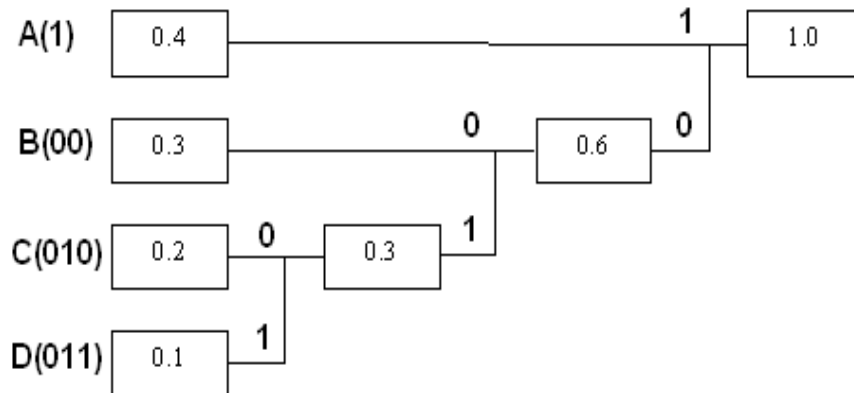
$$\bar{R} = \sum_{i=0}^n p(i)l(i)$$

where  $l(i)$  is the length of code word

The entropy is the weighted sum of all symbols with non-zero probability

$$H(A) = \sum_{w_i > 0} w_i h(a_i) = - \sum_{w_i > 0} w_i \log_2 w_i$$

**Steps to design Huffman Tree[2]**  
As shown into the Fig.4



**Fig. 4** Huffman Code Tree

**IV. VHDL**

With Rampant Growth in Technology, Electronic Design Automation (EDA) software has been growing at an incredible rate during the past several years[5]. The Hardware description language (HDL) is used to describe hardware for the purpose of simulation, modeling, testing design and verification of digital systems. The most popular are VHDL

- 1.Sort source outputs in decreasing order of their probabilities.
- 2.Merge the two least-probable outputs into a single output whose probability is the sum of the corresponding probabilities.
- 3.If the number of remaining output is 2, then go to the next step, otherwise go to step1
- 4.Arbitrarily assign 0 and 1 as code words for the two remaining outputs.
- 5.If an output is the result of the merger of two outputs in preceding step, append the current code word with a 0 and 1 to obtain the code word for the preceding outputs and then repeat the step. If no output is preceded by another output in a preceding step then stop.

The final Coding for Four different symbols A,B,C,D is as shown into **Table.1**

Symbol	Code
A	1
B	00
C	010
D	011

**Table 1:** Result of Huffman Coding

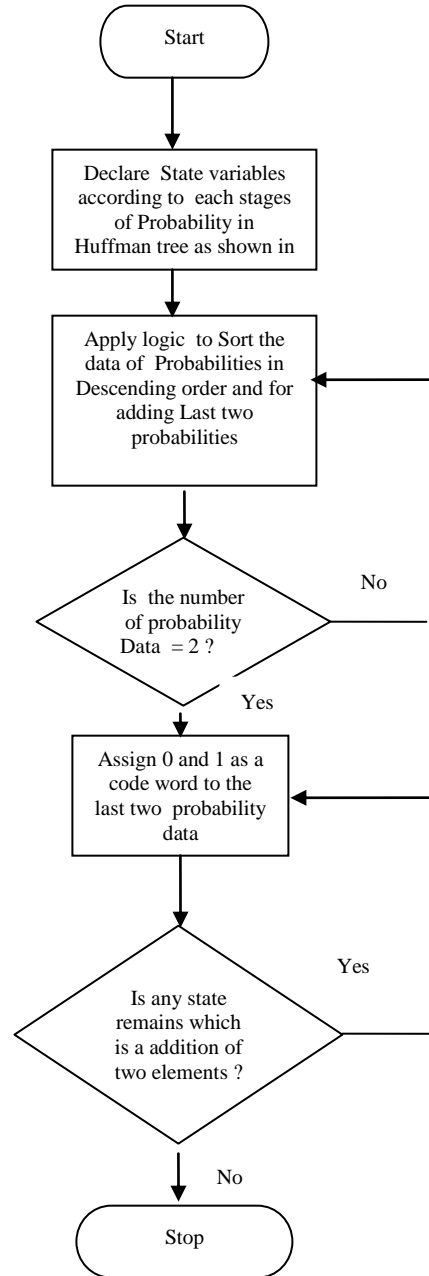
(Very high speed integrated circuit Hardware description language) and Verilog. The most favorable type of HDL used today is VHDL. VHDL is acronym for VHSIC HDL (Very High speed Integrated circuit Hardware Description language), in turn refers to the Very High speed Integrated Circuit Program . This program was sponsored by Department of defense ( DoD) with the goal of developing a new generation of High-Speed Integrated circuit[4]. In VHDL design, a

designer will code the digital design in terms of VHDL code as opposed to the conventional method of Schematic capture. Then this code can be synthesized using VHDL Synthesis tool. VHDL has totally different sets of statements for the sequential and concurrent parts of the program. There are three types of data objects in VHDL namely Variables, Constants, and Signals that hold same data type but have different properties.

**Capabilities of VHDL[6]**

1. The language can be used as an exchange medium between chip vendors and CAD tool users.
  2. The language can be used as a communication medium between different CAD and CAE.
  3. The language is not technology specific but is capable of supporting technology specific features.
  4. Various digital modeling techniques, such as finite state machine descriptions, algorithmic descriptions and Boolean equations, can be modeled using this language.
  5. Language is an IEEE and ANSI standard; therefore, models described using this language is portable.
  6. The language supports three basic programming styles, Structural, Data Flow, Behavioral
- Arbitrarily, large designs can be modeled using the language, and there are no limitations imposed by the language on the size of design

V. SOFTWARE FLOW IN VHDL



**Fig.5** Flow chart

VI. CONCLUSIONS

Entire Laboratory standard model of Huffman coder is implemented on FPGA (Xilinx XC3S400 PQ208) Prototype board using VHDL. Output codes for different symbols are checked using Light emitting diodes available on FPGA prototype board. FPGA based Huffman coder can be utilized in Many data communication systems like video transmission, Audio transmission.etc...FPGA is High speed Integrated circuit , consume less power and small in size, So

FPGA is suitable for Designing a large communication system. As a future expansion one can synthesized entire Embedded communication system including Source coder, Modulator, Demodulator and Source decoder inside FPGA using Hardware description Language.

#### ACKNOWLEDGEMENT

Authors would like to acknowledge support provided by EC Department, A.D.Patel Institute of Technology, New Vallabh Vidyanagar, Gujarat, India where the work has been carried out.

#### REFERENCES

- [1] D.A.Huffman, A method for the construction of minimum redundancy codes, In Proc.Inst.Radio Eng,Vol.40, No.9, PP 1098-1101, Sep-1952
- [2] John G. Proakis, Mosoud Salehi, Communication System Engineering, Pearson Education, Second Edition, PP 276-279
- [3] Hwei, Hsu, Analog and Communications, Tata McgrawHill Edition, Second Edition, PP-255-257
- [4] Sudhakar Yamanchili, Introductory VHDL from Symulation to Synthesis, Person Education, PP-1-2, PP 36-38
- [5] Weng Fook Lee, VHDL Coding and Logic with Synthesis with Synopsys, Academic Press,Tokyo, PP-3-4
- [6] CMOS Digital Integrated Circuits by Sung Mo Kang, Yusuf Leblebici, TMH, PP 28-29
- [7] J.Bhaskar, VHDL Primer, Pearson Education, third edition , PP- 2-6
- [8] R.C.Joshi,, P.M.Koradia FPGA based implementation of PN sequence generator December-2006 Volu-30
- [9] V.K.Thakar, R.C. Joshi, P.M.Koradia, FPGA based implementation of Embedded Entry-Exit Register for Automation Applications, Proceddings of National Conference NCECcube-2007, Pune, India
- [10] Huffman Coding , WikiPedia
- [11] “Mechatronics Made Laboratory Manual”



# Medium Access Control Protocol for Energy Efficiency in Wireless Sensor Networks

A. Prashant J Bagga, B. Nagendra P Gajjar

**Abstract--**Wireless sensor networks provide wide variety of real time application. It can collect and process vast amount of data from environment like pollution, traffic conditions, weather, industrial process monitoring, and condition based maintenance. But due to lower sensing range of these networks, dense networks are required, which bring the necessity to achieve an efficient medium access (MAC) protocol subject to power constraints. Various MAC protocols with different objectives were proposed for wireless sensor networks. In this paper, we have proposed– MAC protocol for demonstrating saving in the energy consumption from all the sources of energy waste like idle listening, collision, overhearing and control overhead

**Index Terms—** Environment monitoring, idle listening, Latency analysis, sleep-wake up cycle.. MAC

## I. INTRODUCTION

Researchers in the Life Sciences are becoming increasingly concerned about the potential impacts of human presence in monitoring plants and animals in field conditions. At worst it is possible that chronic human disturbance may distort results by changing behavioral patterns or distributions.

These considerations are of particular importance for studying the bird behavior like Nalsarovar bird sanctuary located in Gujarat. For purposes of automation of data collection and reduction of human intervention in areas of interest, sensor networks can be considered. Deploying sensor networks will not interfere with existing life. Recent developments in wireless network technology and miniaturization now make it possible to realistically monitor the natural environment. Instrumentation of natural spaces with numerous networked micro-sensors can enable long-term data collection at scales and resolutions that are difficult, if not impossible, to obtain otherwise. The intimate connection with its immediate physical environment allows each sensor to provide localized measurements and detailed information that is hard to obtain through traditional instrumentation. The combination of storage and in-node processing enable them to perform triggering functions suitable for some applications and protocols.

Environmental monitoring is a significant driver for wireless sensor network research, promising dynamic, real-time data about monitored variables of an area and so enabling many new applications. Because of this, almost all real experiments were conducted with this application background. In particular, the first published experience with real deployments of sensor networks were about habitat monitoring [3]. Only recently other application backgrounds

such as wildfire monitoring were considered in real experiments.

Unfortunately, due to the innovative nature of the technology, there are currently very few environmental sensor networks in operation that demonstrate their value. Examples of such networks include NASA/JPL's project in Antarctica [4], Berkeley's habitat modeling at Great Duck Island [1], the CORIE project which studies the Columbian river estuary [5], deserts [6], volcanoes [7] and glaciers [8]. The research efforts in these projects are constantly thriving to a pervasive future in which sensor networks that would expand to a point where information from numerous such networks (e.g. glacier, river, rainfall and oceanic networks) could be aggregated at higher levels to form a picture of the environment at a much higher resolution.

Sensor nodes are the network components that will be sensing and delivering the data. Node transmits its data to its neighbouring nodes or simply passes the data as it is to the Task Manager. Sensor nodes can act as a source or sink in the sensor field. The function of a source is to sense and deliver the desired information. Thus, a source reports the state of the environment. On the other hand, a sink is a node that is interested in some information a sensor in the network might be able to deliver. Gateways allow the scientists and system managers to access nodes through personal computers (PCs), personal digital assistants (PDA) and Internet. In a nutshell, gateways act as a proxy for the sensor network on the Internet.

The Task Manager will connect to the gateways via some media like Internet or satellite. Task Managers comprise of data service and client data browsing and processing. These Task Managers can be visualized as the information retrieval and processing platform. All information (raw, filtered, processed) data coming from sensor nodes is stored in the task managers for analysis. Users can use any display interface (i.e. PDA, computers) to retrieve or analyze data either locally or remotely.

The wireless sensor networks which have sensing, computation and communication functions to move packets from sensor nodes to final servers, consume quantities of energy that must be taken into account to forecast the life cycle of a network and maximize it.

Here an attempt is made to analyze the evaluation of MAC layer protocol for environmental application using simulation tool. S-MAC (Sensor – MAC)[3] uses techniques to reduce energy consumption and support self configuration and also supports low duty cycle operation in a multihop network. Nodes form virtual clusters based on common sleep

schedules to reduce control overhead and enable traffic adaptive wake-up. S-MAC uses in-channel signaling to avoid overhearing unnecessary traffic. Finally S-MAC applies message passing to reduce contention latency for applications that require in-network data processing. The report is organized into following section: section II describes the S-MAC protocol for environmental application, section III discusses the simulation of S-MAC protocol, section IV presents the result and analysis of simulation process and section V discusses the conclusion and future work. The paper ends with references in section VI

## II. S-MAC PROTOCOL

S-MAC tries to reduce energy wastes from all of sources. To reduce control overhead & latency, S-MAC introduces coordinates sleeping among neighbouring nodes. Latency can be important or unimportant depending upon application.

In applications such as surveillance or monitoring, nodes will be vigilant for long time, but largely inactive until something is detected. These applications can often tolerate some additional messaging latency, because the network speed is orders of magnitude faster than the speed of the physical object. S-MAC lets node sleep periodically if they are idle. This design reduces energy consumption but increases latency since sender must wait for the receiver to wake up before it can send data. Another technique, called adaptive listen reduces this latency.

S-MAC re-introduces the concept of message passing to efficiently transmit long messages. Message passing saves energy by reducing control overhead and avoiding overhearing.

Periodic listen and sleep

S-MAC reduces the listen time by putting nodes into periodic sleep state as shown in figure 1 below.

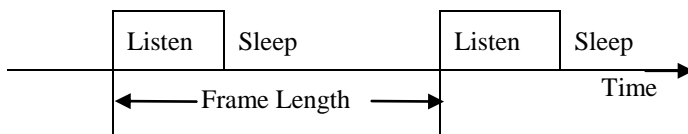


Figure: 1 Periodic Sleep and listen

Neighbouring nodes synchronize their listen / sleep schedule to reduce control overhead. Not all neighbouring nodes can synchronize together in multi-hop network.

Nodes exchange their schedule by periodically broadcasting a SYNC packet to their immediate neighbours. A node can talk to its neighbour at their scheduled listen time, thus ensuring that all neighbouring nodes can communicate even if they have different schedules. The period for a node to send a SYNC packet is called a synchronization period.

Nodes form virtual clusters around common schedules. One advantage of this loose coordination is that it can be more robust to topology change than cluster based approaches. Disadvantage of this scheme is increased latency due to periodic sleeping.

Collision avoidance:

S-MAC follows similar procedures as the 80211 does for collision avoidance, including virtual and physical carrier sense and the RTS/CTS exchanges for hidden terminal problem.

There is a duration field in each transmitted packet that indicates how long the remaining transmission will be. If a node receives a packet destined to another node, it knows how long to keep silent from this field. The node records this value in variable NAV (Network allocation vector) and sets a time for it. Every time when the tuner fires, the node decrements its NAV until it reaches zero. If NAV is not zero, node determines that medium is busy. This is called virtual carrier sense. Physical carrier sense is done at physical layer by listening to channel for possible transmissions. Carrier sense time is randomized within a contention window to avoid collisions & starvations. The medium is determined as free if both virtual and physical carrier sense indicates that it is free.

All senders perform carrier sense before initiating transmission. If a node fails to get a medium it goes to sleep and wakes up when the receiver is free and listening again. Broadcasts packets are send without RTS/CTS and unicast packets follow the sequence of RTS/CTS/DATA/ACK between the sender and receiver. After RTS/CTS, sender and receiver will use their normal sleep time for transmission of data packets. They do not follow sleep schedule until they finish the transmission.

S-MAC effectively addresses energy wastes due to idle listening & collisions.

Advantages of S-MAC:

- Energy waste caused by idle listening is reduced.
- It has simplicity in implementation.
- Overhead of time synchronization is prevented with sleep schedule announcement.

Disadvantages of S-MAC

- Broadcast data packets do not use RTS/CTS which incurs collision probability.
- Adaptive listening incurs overhearing or idle listening if packets are not destined to the listening node.
- Sleep & listen periods are predefined and constants, which decreases the efficiency of the algorithm under variable traffic load.

## III. SIMULATION OF MAC PROTOCOL FOR ENVIRONMENTAL APPLICATION

The environmental application requires continuous sampling of data at defined rate. There are two types of data: sampling and triggered. Sampling data is obtained by sampling a certain parameter a given number of times every day while triggered data is disseminated after a certain event has happened. For energy saving purposes, it is important to differentiate between these two types of data. The S-MAC protocol is proposed to exploit the advantages that sampling data has from an energy saving perspective and, at the same time, cope with latency requirements of triggered data.

Sampling data has two great advantages: first, the number of samples to take in a given period of time is known in advance and second, instants to take the samples are also known. This fact leads us to the idea that between two consecutive sample instants, the communication functions of two nodes is almost null. In this way, the S-MAC protocol exploits this fact to save energy by turning off its radio between two consecutive sample instants for data transmissions. Significant energy savings can be achieved by this operation as idle listening is the most energy consuming operation. However, if the radio is simply turned off, no triggered packets can be transmitted from originating nodes to the base station in a reasonable time. In such situation, triggered packets would be queued up and would also wait for the next available active time slot to be transmitted; what would create a long delay for triggered data, which would ideally have to be transmitted without delay. Furthermore, collisions would increase dramatically because all nodes in the network would content for the medium when the next time slot started. The S-MAC uses RTS/CTS/DATA/ACK mechanism for exchange of packets between nodes protocol.

Here, S-MAC protocol is simulated using Ns-2 to study the behaviour of the protocol for suitability of environmental application. Following are the assumptions made for energy analysis:

- Sampled packets are small enough to be transmitted in a single listen interval.
- Only one node in the network generates sampled packets.
- There is a single route to Base Station.
- Each node has only two neighbors.
- There are no collisions.
- There are no retransmissions.

Network configuration shown in Figure 1 has the following characteristics:

- The four nodes (0, 1, 2, 3) are on a straight line with 150m in apart.
- Node 0 can reach only node 1, 1 can reach 0 and 2, 2 can reach 1 and 3 and 3 only 2.
- The objective of each node is to transmit its data packets to node 3 (Base Station).
- The synchronization and control information is also exchanged between neighbors.

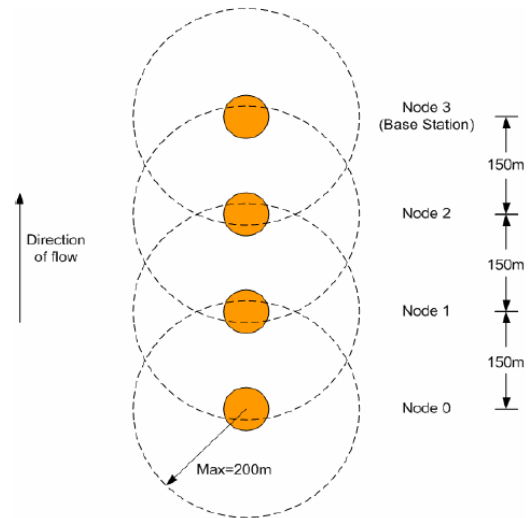


Figure: 1 Initial Node set as visualized in Network

Due to the number of nodes used in the simulation, all nodes must have the same listen/sleep schedule, forming a single virtual cluster. At 100 second, node 0 starts sending 10-byte data packets with a Exponential traffic generator at a mean sending rate. This rate of transmission can also be expressed as the mean time between consecutive packets that we call message inter-arrival time. Each simulation is run at constant duty cycle. For each given constant rate the duty cycle is changed from 10% to 50%. At the end of each simulation, the remaining energy a node is saved for further computations. Then, to compute the total energy consumed in each simulation, the remaining energy is subtracted from initial energy configured to get the energy consumed in a node.

IV. RESULTS AND ANALYSIS

Latency Analysis

Average delay per packet is calculated as:

$$\text{Average delay per packet} = \frac{\text{Total end to end delay for all received packets}}{\text{Total packets received}}$$

TABLE 1

Average delays (ms) per packet for range of duty cycles

% Duty Cycle	End to end delay	Duty cycle	End to end delay
20	864	60	382
30	645	70	303
40	544	80	288
50	435	90	229

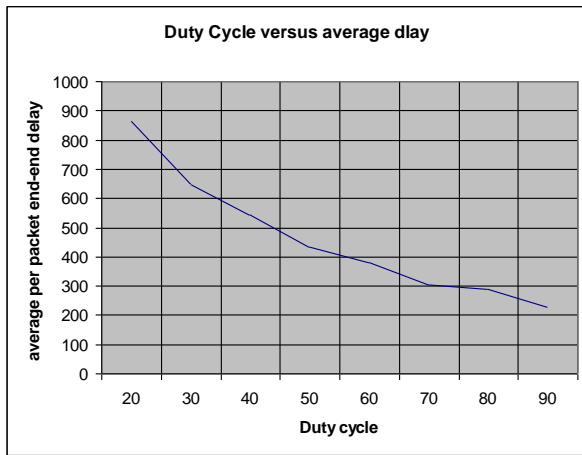


Figure: 2 End-to-end delays for each packet for various duty cycles

Following observation can be made from the above figure 2: As duty cycle is increased the average end-to-end delay time of a packet decrease.

At lower duty cycle, the energy consumed is less but the latency is increased.

In our application of environmental monitoring, time criticality of data is not important. There is no real time quality of service requirements. Hence in less time critical applications like environmental applications, the S-MAC gives more energy savings at the cost of increased latency. The applications like surveillance system or disaster management system, where real time availability of data is critical, the S-MAC with fixed duty cycle will prove to be less effective.

**Energy Analysis**

The results obtained are organized into Tables 2 to 4. In such tables, for each message inter-arrival time, five simulations results are listed. The message inter-arrival period can help us calculate the mean sending rate at which node 0 sends its 10-byte packets:

TABLE: 2  
ENERGY CONSUMED (mJ) IN EACH NODE AT THE END OF SIMULATION USING THE MAC PROTOCOL AT INTER-ARRIVAL TIME OF 100

Energy consumed (mJ) in each node with S-MAC protocol					
Message inter-arrival time Node 0 = 100 ms	Duty Cycle = 10%	Duty Cycle = 20%	Duty Cycle = 30%	Duty Cycle = 40%	Duty Cycle = 50%
Node 0	8104	9384	9483	9562	9574
Node 1	7856	9233	9348	9424	9492
Node 2	7966	9275	9381	9444	9510
Node 3	7897	9205	9357	9407	9482
Total	31823	37097	37569	37837	38058

TABLE: 3  
ENERGY CONSUMED (mJ) IN EACH NODE AT THE END OF SIMULATION USING THE MAC PROTOCOL FOR INTER-ARRIVAL TIME OF 200

Energy consumed (mJ) in each node with S-MAC protocol					
Message inter-arrival time Node 0 = 200 ms	Duty Cycle = 10%	Duty Cycle = 20%	Duty Cycle = 30%	Duty Cycle = 40%	Duty Cycle = 50%
Node 0	6305	9350	9462	9592	9653
Node 1	6022	9187	9339	9450	9537
Node 2	6173	9224	9371	9470	9563
Node 3	6185	9178	9326	9443	9533
Total	24685	36939	37498	37955	38286

TABLE: 4  
ENERGY CONSUMED (mJ) IN EACH NODE AT THE END OF SIMULATION USING THE MAC PROTOCOL AT INTER-ARRIVAL TIME OF 300

Energy consumed (mJ) in each node with S-MAC protocol					
Message inter-arrival time Node 0 = 300 ms	Duty Cycle = 10%	Duty Cycle = 20%	Duty Cycle = 30%	Duty Cycle = 40%	Duty Cycle = 50%
Node 0	4376	9303	9497	9537	9664
Node 1	4234	9102	9359	9455	9572
Node 2	4296	9135	9368	9471	9577
Node 3	4405	9105	9331	9438	9570
Total	17311	36645	37555	37901	38383

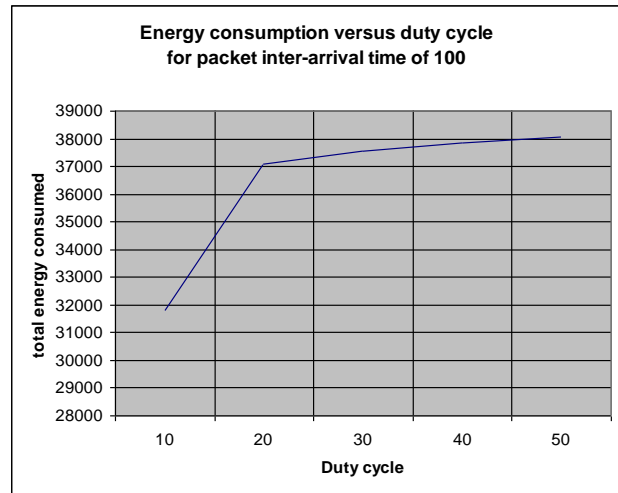


Figure: 3 Energy consumed for inter-arrival time of 100

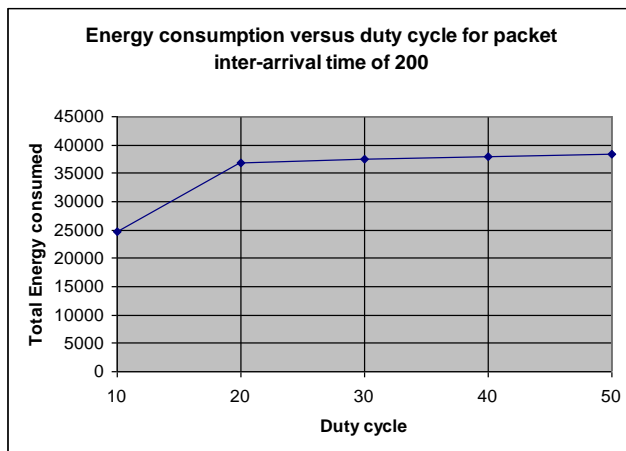


Figure: 4 Energy consumed for inter-arrival time of 200

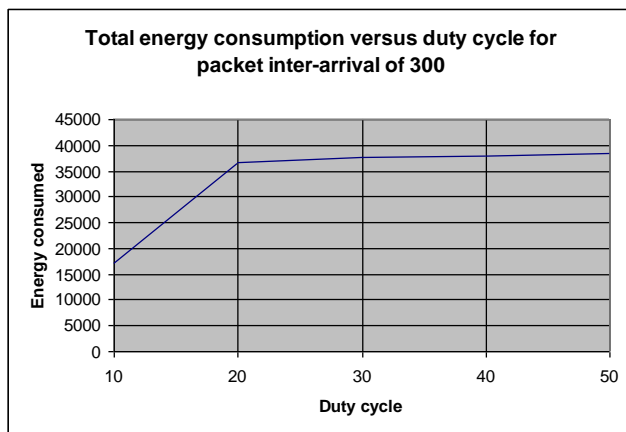


Figure: 5 Energy consumed for inter-arrival time of 300

Following observations are made from above figures 3 to 5:

Energy consumed at low duty cycle is less, as compared to higher duty cycle. This is due to the fact that, the radio is in sleep mode most of the time which reduces energy consumed in idle listening.

The energy consumption is increased as the duty cycle is increased. This is due to the fact that the packet size (10 bytes) is kept small enough; so that it can be send in one cycle time. Hence, if the duty cycle is increased, the sleep time will decrease. This will cause the idle listening by the radio. The idle listening consumes approximately same power as transmitting or receiving. In the simulation setup, the idle power, receive power and transmission power are kept to same 1.0 unit. The fixed duty cycle for S-MAC protocol has a drawback. This calls for adaptive duty cycle, which can adapt to the changes in traffic scenario.

It is also observed that as the message inter-arrival time is increased, bit rate is reduced. This does not effect the total energy consumption by the system, with changes in duty cycle.

## V. CONCLUSIONS AND FUTURE WORK

### Conclusion

As the experimental evaluation of protocols for wireless sensor networks is a costly affair, simulating the protocols for wireless sensor networks with available simulator can provide a better option for to study different aspects.

Network simulator, Ns-2 is a fairly good candidate to simulate wireless sensor networks. Ns-2 comes with rich set of library functions to simulate the protocols for wireless sensor networks. As a case study, S-MAC protocol was simulated using NS-2 simulator to correlate the theoretical background with the simulation results. S- MAC protocol for wireless sensor networks can be used to gather data for wide-area large scale environmental monitoring application. The scheme saves energy by organizing the networks usage changing the running synchronization. Specifically, the proposed protocol uses, a sleep/listen schedule running in top of a previously negotiated one. It can be also concluded that a MAC protocol can be more efficient if it has some information available in Network layer such as number of hops to the Base Station, data arrival rate, etc so that nodes wake up only when a sample is to be taken. This schedule saves more energy by avoiding idle listening. According to simulation results, the proposed scheme is observed to perform better in terms of achievable network lifetime with low duty cycle for the proposed application.

Developing the protocols on sensor nodes with sensor specific network platform and evaluate its performance through simulation and real experiments can provide better understating of this new technology

### Future Work

The documentation for simulating the wireless sensor networks is less explored area. This difficulty was felt when trying to simulate algorithms for wireless sensor networks. Except for direct diffusion, there is no specific documentation available to develop simulation program for wireless sensor networks

The network simulator Ns-2 can be documented to simulate various protocols available for wireless sensor networks.

In future, developing a mechanism to wake-up nodes when a node has the urgency to transmit a triggered packet, which met the requirements of low latency, can be simulated. The experimental evaluation of these protocols for the proposed environmental application can help to validate the simulation results.

## REFERENCES

- [1] Medium Access Control With Coordinated Adaptive Sleeping for Wireless Sensor Networks Wei Ye, Member, IEEE, John Heidemann, Member, IEEE, and Deborah IEEE/ACM transactions on networking, vol. 12, no. 3, June 2004 ah Estrin, Fellow, IEEE
- [2] MEMS Come to Oz Wine Industry", Electronic News June 2004
- [3] A. Cerpa, J. Elson, M. Hamilton, J. Zhao, Habitat monitoring: application driver for wireless communications technology, ACM SIGCOMM'2000, Costa Rica, April '01
- [4] K.A. Delin, R.P. Harvey, N.A. Chabot, S.P. Jackson, Mike Adams, D.W. Johnson, and J.T. Britton, "Sensor Web in Antarctica:

Developing an Intelligent, Autonomous Platform for Locating Biological Flourishes in Cryogenic Environments,” 34th Lunar and Planetary Science Conference, 2003.

- [5] D.C. Steere, et al., “Research Challenges in Environmental Observations and Forecasting Systems,” Proc. ACM/IEEE Int. Conf. Mobile Computing and Networking (MOBICOMM), 2000, pp. 292-299.
- [6] K.A. Delin, S.P. Jackson, D.W. Johnson, S.C. Burleigh, R.R. Woodrow, M. McAuley, J.T. Britton, J.M. Dohm, T.P.A. Ferré, Felipe Ip, D.F. Rucker, and V.R. Baker, “Sensor Web for Spatio-Temporal Monitoring of a Hydrological Environmental,” 35th Lunar and Planetary Science Conference, League City, TX, 2004.
- [7] K. Lorincz, D. Malan, Thaddeus R. F. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnyder, G. Mainland, S. Moulton, and M. Welsh, “Sensor Networks for Emergency Response: Challenges and Opportunities”, Special Issue on Pervasive
- [8] Martinez, K., Hart, J.K., Ong. R. (2004). Environmental Sensor Networks. *Computer*, 37 (8), 50-56.

# An Analysis to Determine the Best Adder Circuit for High Speed Memory BIST

A. Prof. Mayank Kapadiya, B. Prof. Gunjan Jani, C. Dr. N M Devashrayee

**Abstract --** The saying goes that if you can count, you can control. Addition is a fundamental operation for any digital system, digital signal processing or control system. A fast and accurate operation of a digital system is greatly influenced by the performance of the resident adders. Adders are also very important component in digital systems because of their extensive use in other basic digital operations such as subtraction, multiplication and division. Hence, improving performance of the digital adder would greatly advance the execution of binary operations inside a circuit comprised of such blocks. The performance of a digital circuit block is gauged by analyzing its power dissipation, layout area and its operating speed. Here we analyze the fastest adder circuit after the implementation of adder circuitry in BIST. So we can analyze which adder is the fastest at different range of frequency.

**Index Terms – Built In Self Test (BIST).**

## I. INTRODUCTION

Due to the rapid progress in the very large scale integrated (VLSI) technology, an increasing number of transistors can be fabricated onto a single silicon die. For example, a state-of-the-art 130 nm complementary metal-oxide semiconductor (CMOS) process technology can have up to eight metal layers, poly gate lengths as small as 80nm and silicon densities of 200K-300K gates/mm<sup>2</sup>. However, although million-gate integrated circuits (ICs) can be manufactured, the increased chip complexity requires robust and sophisticated test methods.

- A. Assistant Professor, Gandhinagar Institute of Technology, Gujarat Technological University, (email-mayank.kapadiya@git.org.in)
- B. Assistant Professor, Gandhinagar Institute of Technology, Gujarat Technological University, (email- gunjan.jani@git.org.in)
- C. Associate Professor, Institute of Technology, Nirma University of Science & Technology, (email- nmd@nirmauni.ac.in)

Hence, manufacturing test is becoming an enabling technology that can improve the declining manufacturing yield, as well as control the production cost, which is on the rise due to the escalating volume of test data and testing times.

Therefore reducing the cost of manufacturing test, while improving the test quality required to achieve higher product reliability and manufacturing yield, has already been established as a key task in VLSI design.

With the ever increasing complexity and gate counts of modern devices, a number of testability problems have been encountered. One of the fundamental issues is the complexity and size of the test program required to test these devices. The test program is a necessary way to ensure that the high quality standards demanded by the market are met. In brief, two factors are in favor of smaller test programs. Firstly the smaller the program the fewer the number of test vectors and therefore the faster it can be run. Thus, test time can be reduced.

Secondly, if the program is small, problems with the available memory on testing machines can be avoided. One of the main factors which greatly increases the number of test vectors required to test a device is the use of large embedded memories. Hence, if these memories can test themselves a great reduction in test program size can be achieved.

A BIST block is an off-line verification of the Circuit Under Test (it implements the algorithm to test it), as opposed to an on-line verification or concurrent test.

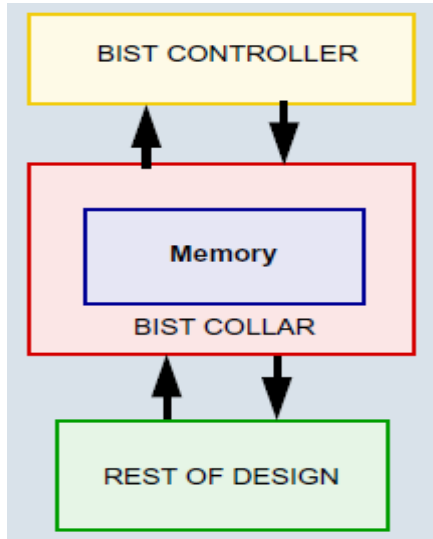


Figure. 1: Block diagram of BIST [1]

Ripple carry adder, Carry look-ahead adder, Carry skip adder and Carry select adder are used in BIST then analyze the critical path timing of the BIST. The analysis is done by using standard ST library in 32nm technology.

The operating condition is as below:

Process	Worst case
Voltage	0.9V
Temperature	125C

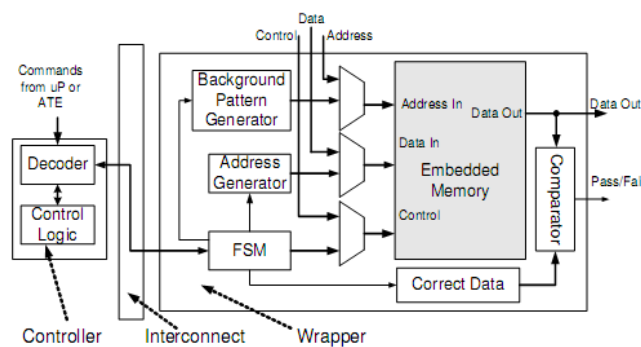


Figure. 2: Generic BIST architecture [2]

Above BIST architecture is used for the following analysis. Typically for the different range of frequency, which one of the adder is fastest and more efficient had been checked. To do the synthesis of the BIST, design compiler of the synopsis is used. [2]

II. IMPLEMENTATION

```

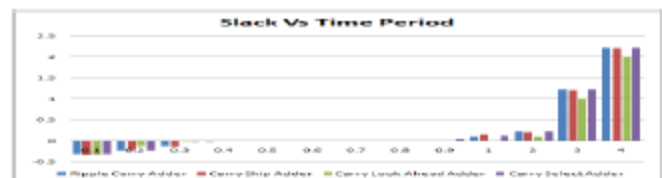
set_max_area 0
Clock Insertion:-
create_clock name clk_m period n
Period is used from n=0.1ns to 1ns using step-size=0.1.
set_clock_uncertainty 0 [get_clocks clk_m]
Input ports:-
set_input_delay 0 i/p pins clock clk_m
Output ports:-
set_output_delay 0 o/p pins clock clk_m
Compile:-
compile_ultra no_boundary_optimization
    
```

Figure.3: Constraints used for the BIST synthesis

Steps required for the analysis of BIST with no boundary optimization:

- a. Use the different adder module in BIST design.
- b. Then do the synthesis for it using above constraints and analyze area and timing requirement for each adder.
- c. Check that the adder is coming in the critical path or not.
- d. If the adder is not coming in the critical path then use set false path [For the path in which adder is not coming] till adder is coming in the critical path and do the synthesis. Analyze the area and timing requirement for adder at different range of frequencies.
- e. Also analyze the area of BIST and timing requirement of critical path in which adder is coming.
- f. Make the table for adders which can show area and timing report for each. So Analyze that which one is the fastest adder among them for different range of frequencies.
- g. Also make graph which can show area and critical path timing comparison of each adder.

III. RESULTS





## IV. CONCLUSION

Comparing the performance metrics for the 8-bit adders implemented in BIST, using Synopsys synthesis tools, the trade off becomes apparent. As can be seen there exist an inverse relationship between time delays, operating speed, and circuit area, in this case the number of CLBs (measure of the area). The ripple carry adder, the most basic of flavors, is at the one extreme of this spectrum with the least amount of CLBs but the highest delay. The carry select adder on the other hand, is at the opposite corner since it has the lowest delay (half that of the ripple carry) but with a larger area required to compensate for this time gain. Finally, the carry look-ahead is middle ground. Power dissipation, for this case study, is in direct proportion to the number of CLBs.

## V. REFERENCES

- [1] M. Abramovici, M. Breuer, and A. Friedman. Digital Systems Testing and Testable Design. IEEE Press IEEE, 1995
- [2] S.M. Al-Harbi and S. K. Gupta. An Efficient Methodology for Generating Optimal and Uniform March Tests. In Proc. IEEE VLSI Test Symposium, pages 231-237, 2001.
- [3] D. Appello, F. Corno, M. Giovinetto, M. Rebaudengo, and M. S. Reorda. AP1500 Compliant BIST-Based Approach to Embedded RAM Diagnosis. In Proc. IEEE Asian Test Symposium, pages 97-102, 2001.
- [4] P. H. Bardell, W. H. McAnney, and J. Savir. Built-In Test for VLSI: Pseudorandom Techniques. John Wiley & Sons, Inc., New York, 1987.
- [5] T. J. Bergfeld, D. Niggemeyer, and E. M. Rudnick. Diagnostic Testing of Embedded Memories Using BIST. In Proc. of the Design, Automation and Test in Europe Conference, pages 305-309
- [6] M. L. Bodoni, A. Benso, S. Chiusano, S. D. Carlo, G. D. Natale, and P. Prinetto. An Elective Distributed BIST Architecture for RAMS. In Proc. IEEE European Test Workshop, pages 119-124, 2000.
- [7] J. T. Chen, J. Rajski, J. Khare, O. Kebichi, and W. Maly. Enabling Embedded Memory Diagnosis via Test Response Compression. In Proc. IEEE VLSI Test Symposium, pages 292-298, 2001.
- [8] H. Cheung and S. K. Gupta. A BIST Methodology for Comprehensive Testing of RAM with Reduced Heat Dissipation. In Proc. IEEE International Test Conference, pages 386-395, 1996.
- [9] R. M. Chou, K. K. Saluja, and V. D. Agrawal. Scheduling Tests for VLSI Systems Under Power Constraints. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 5(2):175-184, June 1997.
- [10] P. T. Gonciari, B. M. Al-Hashimi, and N. Nicolici. Test Data Compression: the System Integrators Perspective. In Proc. of the Design, Automation and Test in Europe Conference, pages 726-731, 2003
- [11] G. L. Craig, C. R. Kime, and K. K. Saluja. Test Scheduling and Control for VLSI Built-In Self-Test. IEEE Transactions on Computers, 37(9):1099-1109, September-1988.
- [12] J. Dreibelbis, J. Barth, H. Kalter, and R. Kho. Processor-Based Built-In Self-Test for Embedded DRAM. Solid-State Circuits, 33(11):1731-1740, November-1998.

# Data acquisition and monitoring parameters of Temperature controllers with MODBUS RTU Protocol using LabVIEW™

A. Sandeep Joshi,

**Abstract--** This paper describes the development of GUI based control software for monitor and data logging of Temperature controllers using LabVIEW™ instrument control software with MODBUS RTU protocol.

**Index-Terms** — LabVIEW™, MODBUS, Protocol, Temperature controller, RS-485 to RS-232, Graphical user interface.

## I. INTRODUCTION

Design graphical user interface for temperature controllers and data logger has been designed in LabVIEW™ environment. Model PID-723 of Libratherm Instruments Microcontroller based PID Temperature Controller with Switching and Analog output (4 to 20mA) is used for temperature measurement. It has input as 'K' type thermocouple for 0 to 1200°C display range. PC interfacing is done by two wire RS 485 MODBUS RTU protocol.

Three PID Controllers are connected in Multi drop network. Two convertors like RS 485 to RS 232 and RS 232 to USB convertor is used with two wire twisted pair cable.

LabVIEW™ programs are designed to facilitate data analysis, as well as offer numerous display options. LabVIEW™ can communicate with hardware like GPIB, PXI, VXI, RS-232 and RS-485 devices.

There are so many advantages of LabVIEW™ like: Graphical user interface, Drag and drop built-in functions, Reduces cost and preserve investment, Flexibility and scalability, Connectivity and instrument control, Compiled language for fast execution, etc.

Currently, the operation of the temperature controller is carried out manually using its keypad. The main motivation of the development of this GUI is for easier user convenience as well as for future use when the laser will be used with the plasma accelerator system, where the temperature monitor will be carried remotely and the temperature system will be integrated with the accelerator system operations *via* LabVIEW™ instrument control

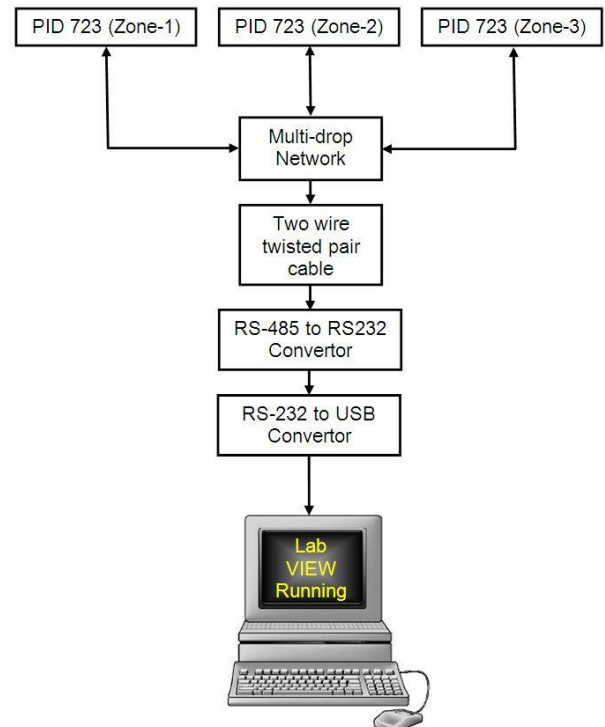


Fig.1.Schematic diagram of the temperature monitor system showing the block of I/O interfaces.

software.

Routine monitoring temperature operation which is currently being carried out manually will be automated using the GUI.

The project requires development of the essential LabVIEW™ virtual instrument modules, GUI interfaces and the required display screens for parameter display and documentation of the operational parameters of the temperature system.

The I/O control of the temperature controller is established via a standard serial interface coupled to the PC through a RS-485 to RS-232 and RS-232-to-USB interface. LabVIEW™ tools like web-publishing Tool, structure palette, string, VISA palette have been used to design this graphical user interface. The temperature control is designed to be controlled either from the control

START	SLAVE ADDRESS	FUNCTIONS CODE	DATA BITS	CRC CHECK	STOP
-------	---------------	----------------	-----------	-----------	------

Fig. 2 Message format of RTU modbus protocol.

panel or from a remote PC but not from both at the same time.

II. MODBUS PROTOCOL IN RTU MODE

Modbus transmission protocol was developed by Gould Modicon (now AEG) for process control systems. In contrast to the many other buses discussed, no interface is defined. The user can therefore choose between RS-422, RS-485 or 20 mA current loops, all of which are suitable for the transmission rates, which the protocol defines. Although the Modbus is relatively slow in comparison with other buses, it has the advantage of wide acceptance among instrument manufacturers and users. About 20 to 30 manufacturers produce equipment with the Modbus protocol and many systems are in industrial operation.

It can therefore be regarded as a *de facto* industrial standard with proven capabilities. A recent survey in the well-known American Control Engineering magazine indicated that over 40% of industrial communication applications use the Modbus protocol for interfacing.

The Modbus protocol provides frames for the transmission of messages between master and slaves. In

, 2 bits if no parity. ” Error Check Field: Cyclic Redundancy check. See in Fig.2.

III. RS-485 TO RS-232 WITH MULTI DROP NETWORK

RS-485 permits a ‘multi drop’ network connection on 2 wires and allows reliable serial data communication for: Distances of up to 1200 m (4000 feet, same as RS-422).Data rates of up to 10 Mbps (same as RS-422).Up to 32 line drivers on the same line. Up to 32 line receivers on the same line. See in Fig.3.

Fig.4 is block diagram of RS-485 to RS-232 converter. Interface converters can also be used to increase the effective distance between two RS-232 devices. RS-232/485 interface converters are very similar and provide bidirectional full-duplex conversion for synchronous or asynchronous transfer between RS-232 and RS-485 ports.

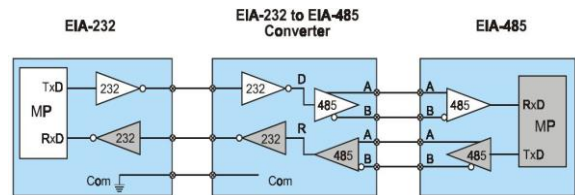


Fig.4 Internal diagram of RS-485 to RS-232 converter.

IV. TEMPERATURE MONITOR AND DATA LOGGING MODULE

The block diagram of this module is seeing in Fig.5.

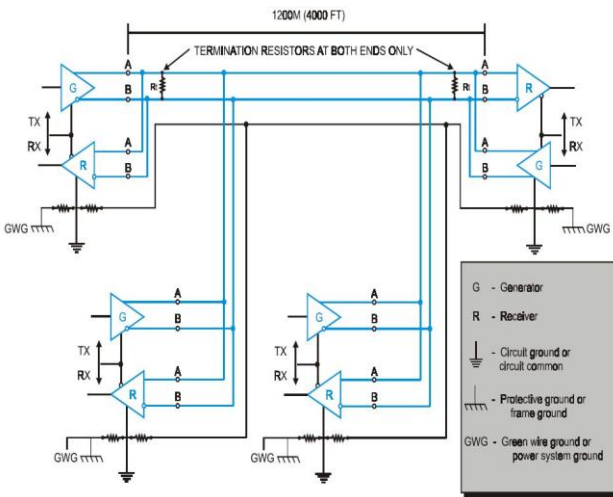


Fig. 3 2-wire multi-drop network.

RTU mode, each byte in a message contains two 4-bit hexadecimal characters. The greater density allows for better throughput ASCII mode at the same baud rate. The message is below:” Coding system: 8-bit binary, hexadecimal 0 to 9, A to F. Two hexadecimal characters are contained in each 8-bit field of the message. ” Bits per bytes:1 start bit,8 data bits, LSB sent first, 1 bit for odd or even parity, no bits for no parity,1 stop bit if parity is used

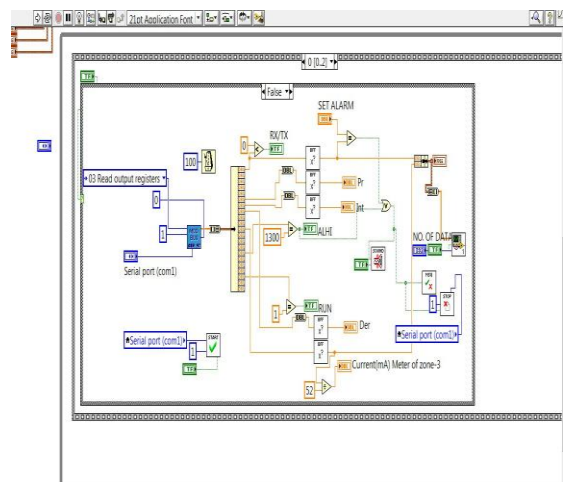


Fig.5 Block diagram of Monitor and data acquisition of PID temperature controller

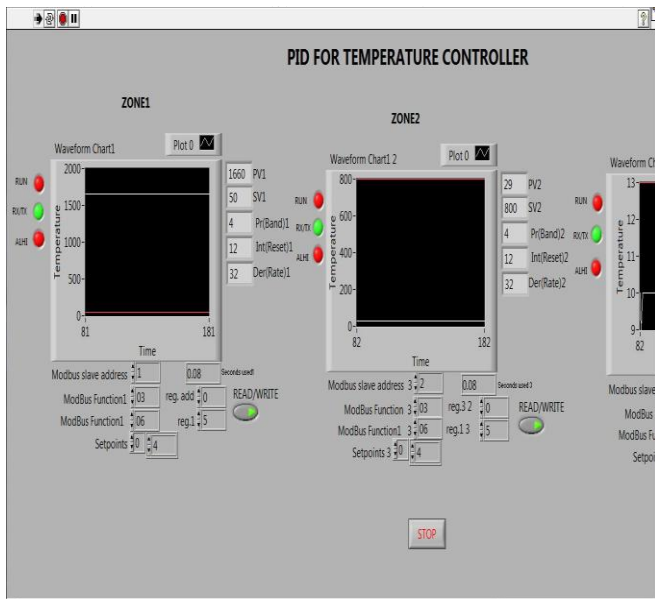


Fig.6. Front-panel for PID controller zone -1 and zone-2.

Finally, modules of temperature monitor module are shown in Fig.7.

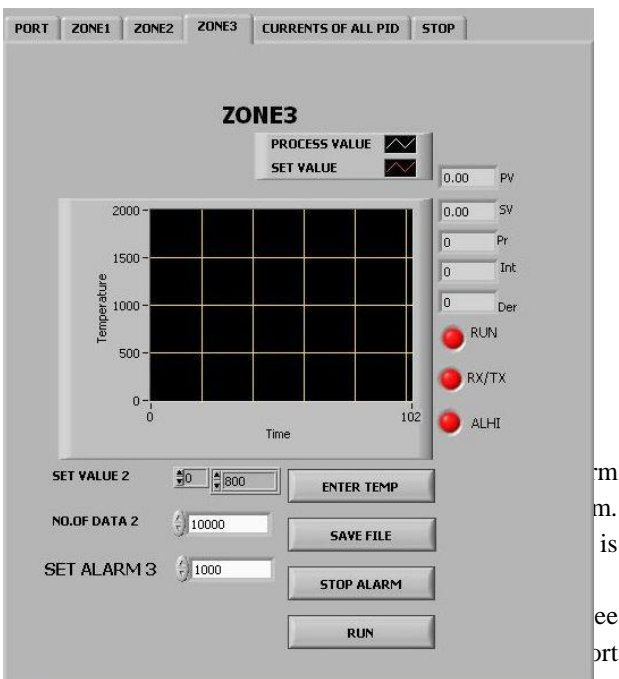


Fig.7. Front-panel for GUI for Temperature monitors and data

Stop tab has 'Stop' button by which user can stop execution.

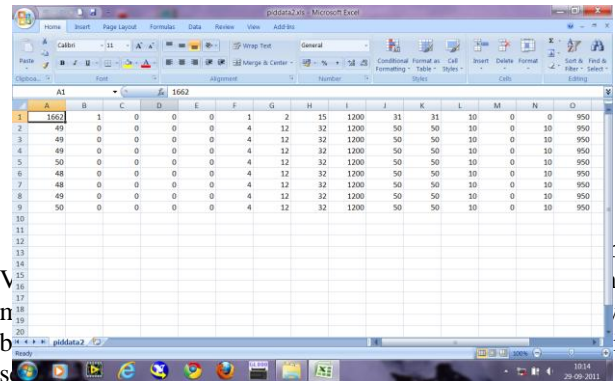
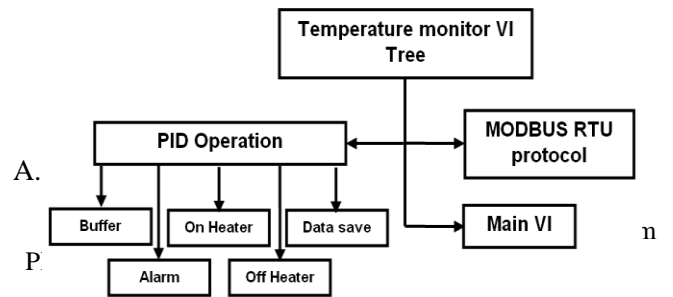


Fig.8 Data acquire in excel file for zone-2.



A. *Alarm command*  
It generates the sound and stop heater at temperature which user has set. It is only safety purpose.

C. *On heater command*  
Heater is on when this sub VI will be run. Parameters of PID controllers cannot change when heater will be on.

D. *Off heater command*

Heater is off when this sub VI will be run. Parameters can change when heater will be off.

E. *MODBUS RTU Protocol*

All parameters of PID controllers are read in read buffer indicator in LabVIEW™ front panel in hex format when this sub VI will be run. Enter the slave addresses, baud rate, and function codes before run VI.

## VI. CONCLUSION

The user interface for monitoring parameters of PID temperature controller was designed in LabVIEW™ software. The GUI of PID controllers are implemented with two wire RS-485 using RTU mode MODBUS protocol using single RS-485 to RS-232 convertor. The data of these parameters were store in the excel file. The data of parameters were store for individual PID temperature controller faster than connected PID temperature controllers in multi-drop network. In present system data acquisition was done by stand alone model PID-723 using LabVIEW™. In low cost solution, one can develop on hardware having signal conditioning and one can design own PID controller in LabVIEW™. All operation will be done on internet in LabVIEW™ environment if all computers are connected in LAN.

## ACKNOWLEDGMENT

I would like to thank Dr. Ravi A.V. Kumar and Mr.Ankit Sharma for all their timely and useful guidance and help which resulted in the successful completion of this project.

I would also like to thank Mr. Yogesh Yeole for helping me overcome many problems that I faced while using LabVIEW™. I would also like to thank Mr. Mohandas, Ms. Kanchan Mahavar, Mr. Abhijeet and Mr. Jinto Thomas of IPR for help rendered on various occasions right through this project work.

## F. REFERENCES

- [1] John Park, Steve Mackay, and Edwin Wright, "Practical Data Communications for Instrumentation and Control," 2003.
- [2] NATIONAL INSTRUMENT LabVIEW™ 2010 –software.
- [3] JOVITHA JEROME, "VIRTUAL INSTRUMENTATION USING Lab VIEW™," Published by Asoke K. Ghosh., 2009.
- [4] Hexin RS-485 to RS-S232 interface converter Model-485.
- [5] Microcontroller based PID Temperature controller Model PID-723 of Libratherum instrument manual.
- [6] Bella G Liptak,"Process Software and Digital networking"<sup>3<sup>rd</sup></sup>

# A Review of Dynamic Models of Induction Motor

A. P. R. Mankad, B. A. R. Chudasama

**Abstract**— An Induction motor is very simple in construction but complex in control. With Vector control method it is possible to control the induction motor like separately excited DC motor. Also recent methods such as Direct Torque Control makes it even simpler to extract high dynamic performance from induction motors comparable to dc motors. But all these methods require dynamic model of induction motor which is very different from steady state equivalent circuit model. A survey of evolution of various dynamic models for induction motors is presented in this paper.

**Keywords**— d-q model, synchronously rotating reference frame, stationary reference frame. Vector control.

## I. INTRODUCTION

For adjustable speed induction motor drives open loop V/f control is so far most simple and widely used method. However, it gives inferior performance due to complex coupled nature of IM. Vector controlled Induction motor drives give decoupled DC machine like control so they are quite often used in high performance industrial applications. There is a strong interest by the drives manufacturers to replace V/f drives by Vector controlled drives, as the technology is growing and there is only a minimum extra cost for the second solution. To make use of Vector control for induction motor drives, dynamic d-q model is required. Conventional steady state equivalent circuit model can not be used because of coupling problem. Coupling of flux and torque makes it impossible to achieve good dynamic response from an induction motor drive.

An induction motor can be looked on as a transformer with secondary changing its position. Here coupling coefficients between the stator and rotor phases continuously change with rotor position  $\theta_r$ . The dynamic performance of an induction motor becomes complex because of three rotor winding rotation with respect to stator and giving rise to time varying mutual fluxes. First attempt was done to represent three phase machine by an equivalent two phase machine as shown in Fig.1, with  $d^s$ - $q^s$  corresponds to stator direct and quadrature axes and  $d^r$ - $q^r$  corresponds to rotor direct and quadrature axes. Although it is somewhat simple, the problem of time-varying parameters still remains. R. H. Park, in the 1920s, proposed

a new theory of electric machine analysis to solve this problem. He formulated a change of variables which, in effect, replaced the variables (voltages, currents, and flux

linkages) associated with the stator windings of a synchronous machine with variables associated with fictitious windings rotating with the rotor at synchronous speed. Essentially, he transformed, or referred, the stator variables to a synchronously rotating reference frame fixed in the rotor. With such a transformation (called Park's transformation), he showed that all the time-varying inductances that occur due to an electric circuit in relative motion and electric circuits with varying magnetic reluctances can be eliminated. Later, in the 1930s, H. C. Stanley showed that time-varying inductances in the voltage equations of an induction machine due to electric circuits in relative motion can be eliminated by transforming the rotor variables to variables associated with fictitious stationary windings. In this case, the rotor variables are transformed to a stationary reference frame fixed on the stator. Later, G. Kron proposed a transformation of both stator and rotor variables to a synchronously rotating reference frame that moves with the rotating magnetic field. D. S. Brereton proposed a transformation of stator variables to a rotating reference frame that is fixed on the rotor. In fact, it was shown later by Krause and Thomas that time-varying inductances can be eliminated by referring the stator and rotor variables to a common reference frame which may rotate at any speed (arbitrary reference frame).

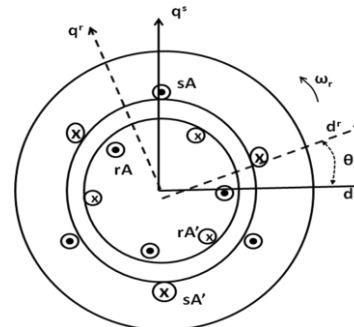


Fig. 1 Stator and rotor windings with relative motion

## II. STEADY EQUIVALENT CIRCUIT MODEL

Fig. 2 shows the approximate equivalent circuit model very popularly used for open loop speed control of induction motor drives. Here  $R_m$  is neglected. All the parameters are referred to stator side. This model is good enough for steady state calculations of various motor parameters with load. But transient disturbances can not be taken care of with this model as it is having inherent coupling of voltage and frequency.

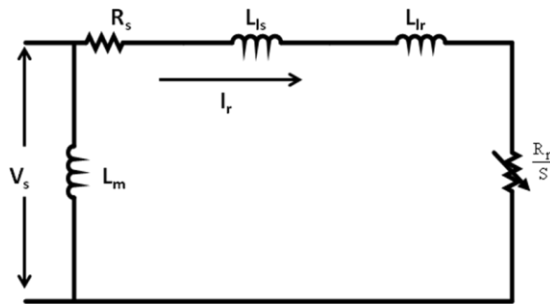


Fig. 2 Approximate equivalent circuit of induction motor

III. THREE PHASE TO TWO PHASE AXES TRANSFORMATION

Consider a symmetrical three-phase induction machine with stationary as-bs-cs axes at 120°-angle apart, as shown in Figure 3. Our goal is to transform the three-phase stationary reference frame (as-bs-cs) variables into two-phase stationary reference frame (d<sup>s</sup>-q<sup>s</sup>) variables and then transform these to synchronously rotating reference frame d<sup>e</sup>-q<sup>e</sup> and vice versa. Assume that the d<sup>s</sup>-q<sup>s</sup> axes are oriented at θ angle, as shown in Figure 3. The voltages V<sub>ds</sub><sup>s</sup> and V<sub>qs</sub><sup>s</sup> can be resolved into V<sub>as</sub>-V<sub>bs</sub>-V<sub>cs</sub> components and can be represented in the matrix form as:

$$\begin{bmatrix} V_{as} \\ V_{bs} \\ V_{cs} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 1 \\ \cos(\theta-120^\circ) & \sin(\theta-120^\circ) & 1 \\ \cos(\theta+120^\circ) & \sin(\theta+120^\circ) & 1 \end{bmatrix} \begin{bmatrix} V_{qs}^s \\ V_{ds}^s \\ V_{os}^s \end{bmatrix} \quad \text{--- (1)}$$

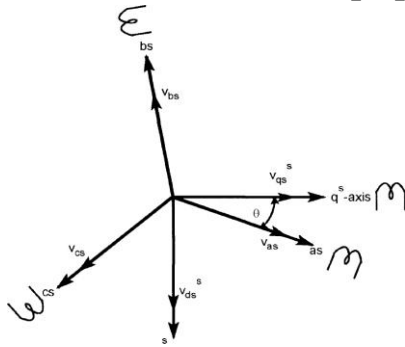


Fig. 3 Three phase to two phase transformation

The corresponding inverse relation is

$$\begin{bmatrix} V_{qs}^s \\ V_{ds}^s \\ V_{os}^s \end{bmatrix} = \frac{2}{3} \begin{bmatrix} \cos\theta & \cos(\theta-120^\circ) & \cos(\theta+120^\circ) \\ \sin\theta & \sin(\theta-120^\circ) & \sin(\theta+120^\circ) \\ 0.5 & 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} V_{as} \\ V_{bs} \\ V_{cs} \end{bmatrix} \quad \text{--- (2)}$$

Where V<sub>os</sub><sup>s</sup> is added as the zero sequence component, which may or may not be present. We have considered voltage as the variable. The current and flux linkages can be transformed by similar equations.

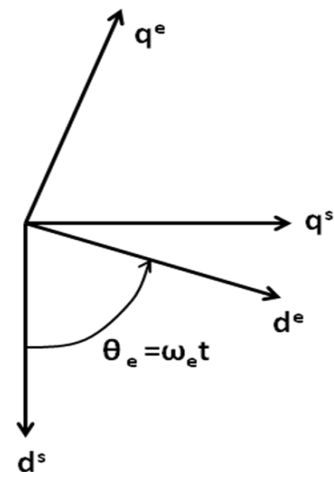


Fig. 4 Stationary d<sup>s</sup>-q<sup>s</sup> and synchronously d<sup>e</sup>-q<sup>e</sup> frames.

Figure 4 shows the synchronously rotating d<sup>e</sup>-q<sup>e</sup> axes, which rotate at synchronous speed ω<sub>e</sub> with respect to the d<sup>s</sup> - q<sup>s</sup> axes and the angle θ<sub>e</sub> = ω<sub>e</sub>t. The two-phase d<sup>s</sup> - q<sup>s</sup> windings are transformed into the hypothetical windings mounted on the d<sup>e</sup> - q<sup>e</sup> axes. The voltages on the d<sup>s</sup> - q<sup>s</sup> axes can be converted (or resolved) into the d<sup>e</sup> - q<sup>e</sup> frame as follows:

$$V_{qs} = V_{qs}^s \cos\theta_e - V_{ds}^s \sin\theta_e \quad (3)$$

$$V_{ds} = V_{qs}^s \sin\theta_e + V_{ds}^s \cos\theta_e \quad (4)$$

Superscript ‘e’ is often not mentioned for parameters referred to synchronously rotation reference frame.

Resolving the rotating frame parameters into a stationary frame, the relations are:

$$V_{qs}^s = V_{qs} \cos\theta_e = V_{ds} \sin\theta_e \quad (5)$$

$$V_{ds}^s = -V_{qs} \sin\theta_e + V_{ds} \cos\theta_e \quad (6)$$

Assume that the three-phase stator voltages are sinusoidal and balanced, and are given by

$$\left. \begin{aligned} V_{as} &= V_m \cos(\omega_e t + \Phi) \\ V_{bs} &= V_m \cos(\omega_e t - \frac{2\pi}{3} + \Phi) \\ V_{cs} &= V_m \cos(\omega_e t + \frac{2\pi}{3} + \Phi) \end{aligned} \right\} \text{--- (7)}$$

Substituting Equation (7) in (5)-(6) yields

$$\left. \begin{aligned} A. \quad V_{qs}^s &= V_m \cos(\omega_e t + \Phi) \\ V_{ds}^s &= -V_m \sin(\omega_e t + \Phi) \end{aligned} \right\} \text{--- (8)}$$

Again substituting Equation (8) in (3) and (4), we get

$$\left. \begin{aligned} V_{qs} &= V_m \cos\Phi \\ V_{ds} &= -V_m \sin\Phi \end{aligned} \right\} \text{--- (9)}$$

Equations (8) and (9) verify that sinusoidal variables in a stationary frame appear as dc quantities in a synchronously rotating reference frame.

The variables in a reference frame can be combined and represented by a complex space vector (or phasor). For example, from Equations (8),

$$\left. \begin{aligned} V_{qs}^s &= V_{qs}^s - jV_{ds}^s \\ &= V_m(\cos(\omega_e t + \Phi) - j \sin(\omega_e t + \Phi)) \\ &= V_m e^{j\theta} e^{j\omega_e t} \\ &= \sqrt{2}V_s e^{j(\theta_e + \Phi)} \end{aligned} \right\} \text{---- (10)}$$

Where  $V_m$  is maximum value and  $V_s$  is rms phasor magnitude. The  $d^e$ - $q^e$  components can also be represented in complex vector form as:

$$\left. \begin{aligned} V_{qs}^e &= V_{qs}^e - jV_{ds}^e \\ &= (V_s^s \cos \theta_e - V_{ds}^s \sin \theta_e) - j(V_s^s \sin \theta_e + V_{ds}^s \cos \theta_e) \\ &= (V_{qs}^e - jV_{ds}^e) e^{-j\theta} \end{aligned} \right\} \text{---- (11)}$$

#### IV. SYNCHRONOUSLY ROTATION REFERENCE FRAME MODEL

Transforming from three phases to two phases and use d-q transformation to stationary reference frame, we can write the following stator circuit equations:

$$V_{qs}^s = R_s i_{qs}^s + \frac{d\psi_{qs}^s}{dt} \quad (12)$$

$$V_{ds}^s = R_s i_{ds}^s + \frac{d\psi_{ds}^s}{dt} \quad (13)$$

Where  $\psi_{qs}^s$  and  $\psi_{ds}^s$  are q-axis and d-axis stator flux linkages, respectively. When these equations are transformed into synchronously rotating  $d^e$ - $q^e$  reference frame, we get following equations:

$$V_{qs}^e = R_s i_{qs}^e + \frac{d\psi_{qs}^e}{dt} + \omega_e \psi_{ds}^e \quad (14)$$

$$V_{ds}^e = R_s i_{ds}^e + \frac{d\psi_{ds}^e}{dt} - \omega_e \psi_{qs}^e \quad (15)$$

Where all the variables are referred to  $d^e$ - $q^e$  reference frame. The last term in Equations (14) and (15) can be defined as speed emf due to rotation of the axes. When  $\omega_e = 0$ , the equations revert to stationary form. If the rotor is not moving, that is,  $\omega_r = 0$ , the rotor equations for a doubly-fed wound-rotor machine will be similar to Equations (14)-(15):

$$V_{qr} = R_r i_{qr} + \frac{d\psi_{qr}}{dt} + \omega_e \psi_{dr} \quad (16)$$

$$V_{dr} = R_r i_{dr} + \frac{d\psi_{dr}}{dt} - \omega_e \psi_{qr} \quad (17)$$

Where all the variables and parameters are referred to the stator. Since the rotor actually moves at speed  $\omega_r$ , the d-q axes fixed on the rotor move at a speed  $\omega_e - \omega_r$  relative to the synchronously rotating frame. Therefore, in  $d^e$ - $q^e$  frame, the rotor equations should be modified as:

$$V_{qr} = R_r i_{qr} + \frac{d\psi_{qr}}{dt} + (\omega_e - \omega_r) \psi_{dr} \quad (18)$$

$$V_{dr} = R_r i_{dr} + \frac{d\psi_{dr}}{dt} - (\omega_e - \omega_r) \psi_{qr} \quad (19)$$

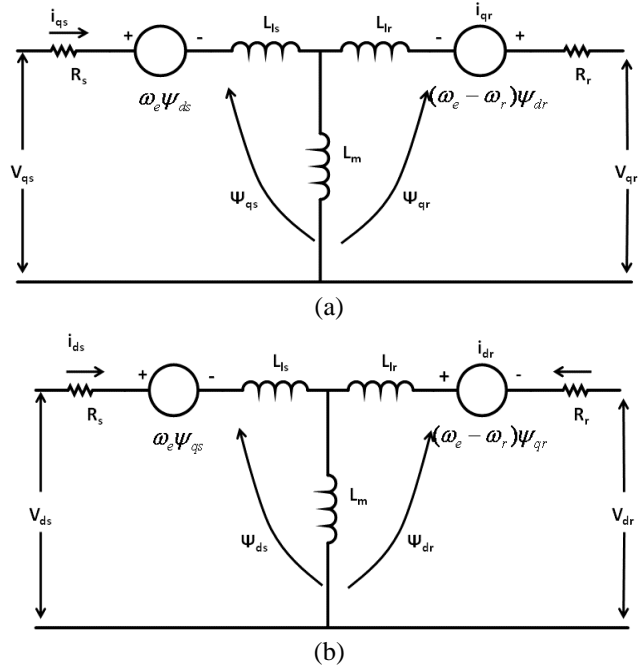


Fig. 5 Dynamic  $d^e$ - $q^e$  model equivalent circuits of machine (a)  $q^e$ -axis circuit, (b)  $d^e$ -axis circuit,

Figure 5 shows the  $d^e$ - $q^e$  dynamic model equivalent circuits that satisfy Equations (14)-(15) and (18)-(19). A special advantage of the  $d^e$ - $q^e$  dynamic model of the machine is that all the sinusoidal variables in stationary frame appear as dc quantities in synchronous frame, as discussed before.

The flux linkage expressions in terms of the currents can be written from Figure 5 as follows:

$$\psi_{qs} = L_{ls} i_{qs} + L_m (i_{qs} + i_{qr}) \quad (20)$$

$$\psi_{qr} = L_{lr} i_{qr} + L_m (i_{qs} + i_{qr}) \quad (21)$$

$$\psi_{qm} = L_m (i_{qs} + i_{qr}) \quad (22)$$

$$\psi_{ds} = L_{ls} i_{ds} + L_m (i_{ds} + i_{dr}) \quad (23)$$

$$\psi_{dr} = L_{lr} i_{dr} + L_m (i_{ds} + i_{dr}) \quad (24)$$

$$\psi_{dm} = L_m (i_{ds} + i_{dr}) \quad (25)$$

Combining the above expressions with Equations (14), (15), (18), and (19), the electrical transient model in terms of voltages and currents can be given in matrix form as

$$\begin{bmatrix} V_{qs} \\ V_{ds} \\ V_{qr} \\ V_{dr} \end{bmatrix} = \begin{bmatrix} R_s + SL_s & \omega_e L_s & SL_m & \omega_e L_m \\ -\omega_e L_s & R_s + SL_s & -\omega_e L_m & SL_m \\ SL_m & (\omega_e - \omega_r) L_m & R_r + SL_r & (\omega_e - \omega_r) L_r \\ -(\omega_e - \omega_r) L_m & SL_m & -(\omega_e - \omega_r) L_r & R_r + SL_r \end{bmatrix} \begin{bmatrix} i_{qs} \\ i_{ds} \\ i_{qr} \\ i_{dr} \end{bmatrix} \text{----(26)}$$

where  $S$  is the Laplace operator. For a singly-fed machine, such as a cage moto  $V_{qr} = V_{dr} = 0$ .



If the speed  $\omega_r$  is considered constant (infinite inertia load), the electrical dynamics of the machine are given by a fourth-order linear system. Then, knowing the inputs  $V_{qs}$ ,  $V_{ds}$ , and  $\omega_r$ , the currents  $i_{qs}$ ,  $i_{ds}$ ,  $i_{qr}$  and  $i_{dr}$  can be solved from Equation (26).

For variable speed the motor load torque dynamic equation which can be used is given as:

$$T_e = T_L + J \frac{d\omega_m}{dt} = T_L + \frac{2}{P} J \frac{d\omega_r}{dt} \quad (27)$$

where  $T_L$ = load torque,  $J$  = rotor inertia, and  $\omega_m$  = mechanical speed. So to evaluate  $\omega_r$  we need to find out  $T_e$  in terms of  $d^e$ - $q^e$  flux linkages and currents. The torque equation in vector form is given as:

$$T_e = \frac{3}{2} \left( \frac{P}{2} \right) \bar{\psi}_m \times \bar{I}_r \quad (28)$$

$$= \frac{3}{2} \left( \frac{P}{2} \right) (\psi_{dm} i_{qr} - \psi_{dm} i_{dr}) \quad (29)$$

or

$$= \frac{3}{2} \left( \frac{P}{2} \right) (\psi_{dm} i_{qs} - \psi_{dm} i_{ds}) \quad (30)$$

Equations (26) to (30) give the complete model of electro-mechanical dynamics of induction motor in synchronously rotating reference frame.

### V. STATIONARY REFERENCE FRAME MODEL

The dynamic machine model in stationary frame can be derived simply by substituting  $\omega_e = 0$  in Equation (26) or in (14), (15), (18), and (19). The corresponding stationary frame equations are given as

$$V_{qs}^s = R_s i_{qs}^s + \frac{d\psi_{qs}^s}{dt} \quad (31)$$

$$V_{ds}^s = R_s i_{ds}^s + \frac{d\psi_{ds}^s}{dt} \quad (32)$$

$$0 = R_r i_{qr}^s + \frac{d\psi_{qr}^s}{dt} - \omega_r \psi_{dr}^s \quad (33)$$

$$0 = R_r i_{dr}^s + \frac{d\psi_{dr}^s}{dt} + \omega_r \psi_{qr}^s \quad (34)$$

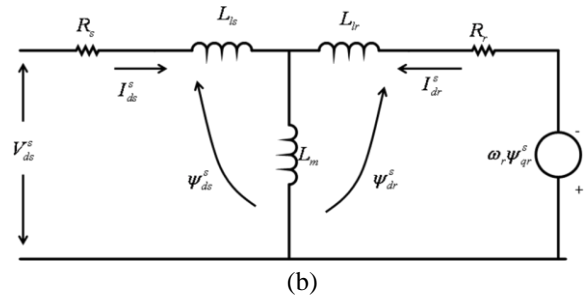
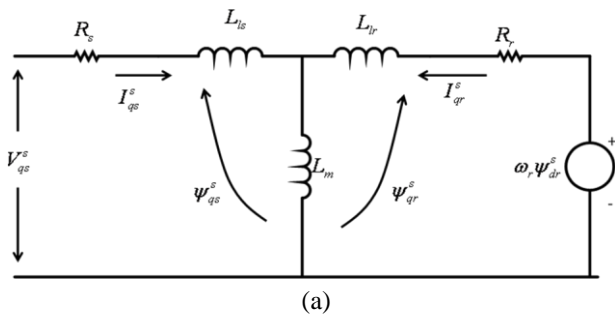


Fig. 6 Dynamic  $d^s$ - $q^s$  model equivalent circuits of machine (b)  $q^s$ -axis circuit, (a)  $d^s$ -axis circuit

Where  $V_{qr} = V_{dr} = 0$ . Figure 6 shows the corresponding equivalent circuits.

The torque Equations (28)-(30) can also be written with the corresponding variables in stationary frame as

$$T_e = \frac{3}{2} \left( \frac{P}{2} \right) (\psi_{dm}^s i_{qr}^s - \psi_{qm}^s i_{dr}^s) \quad (35)$$

$$= \frac{3}{2} \left( \frac{P}{2} \right) (\psi_{dm}^s i_{qs}^s - \psi_{qm}^s i_{ds}^s) \quad (36)$$

or

$$= \frac{3}{2} \left( \frac{P}{2} \right) (\psi_{ds}^s i_{qs}^s - \psi_{qs}^s i_{ds}^s) \quad (37)$$

$$= \frac{3}{2} \left( \frac{P}{2} \right) L_m (i_{qs}^s i_{dr}^s - i_{ds}^s i_{qr}^s) \quad (38)$$

### VI. STATE SPACE REPRESENTATION OF DYNAMIC MODEL

The dynamic machine model in state-space form is important for transient analysis, particularly for computer simulation study. Although the rotating frame model is generally preferred, the stationary frame model can also be used. The electrical variables in the model can be chosen as fluxes, currents, or a mixture of both. Let the flux linkage variables be defined as follows:

$$F_{qs} = \omega_b \psi_{qs} \quad (39)$$

$$F_{qr} = \omega_b \psi_{qr} \quad (40)$$

$$F_{ds} = \omega_b \psi_{ds} \quad (41)$$

$$F_{dr} = \omega_b \psi_{dr} \quad (42)$$

where  $\omega_b$  - base frequency of the machine.

Substituting the above relations in Equations (14)-(15) and (18)-(19), we can write

$$V_{qs} = R_s i_{qs} + \frac{1}{\omega_b} \frac{dF_{qs}}{dt} + \frac{\omega_e}{\omega_b} F_{ds} \quad (43)$$

$$V_{ds} = R_s i_{ds} + \frac{1}{\omega_b} \frac{dF_{ds}}{dt} - \frac{\omega_e}{\omega_b} F_{qs} \quad (44)$$

$$0 = R_r i_{qr} + \frac{1}{\omega_b} \frac{dF_{qr}}{dt} + \frac{(\omega_e - \omega_r)}{\omega_b} F_{dr} \quad (45)$$

$$0 = R_r i_{dr} + \frac{1}{\omega_b} \frac{dF_{dr}}{dt} - \frac{(\omega_e - \omega_r)}{\omega_b} F_{qr} \quad (46)$$

Where it is assumed that  $V_{qr} = V_{dr} = 0$ .

Multiplying Equations (20)–(25) by  $\omega_b$  on both sides, the flux linkage expressions can be written as:

$$F_{qs} = \omega_b \Psi_{qs} = X_{ls} i_{qs} + X_m (i_{qs} + i_{qr}) \quad (47)$$

$$F_{qr} = \omega_b \Psi_{qr} = X_{lr} i_{qr} + X_m (i_{qs} + i_{qr}) \quad (48)$$

$$F_{qm} = \omega_b \Psi_{qm} = X_m (i_{qs} + i_{qr}) \quad (49)$$

$$F_{ds} = \omega_b \Psi_{ds} = X_{ls} i_{ds} + X_m (i_{ds} + i_{dr}) \quad (50)$$

$$F_{dr} = \omega_b \Psi_{dr} = X_{lr} i_{dr} + X_m (i_{ds} + i_{dr}) \quad (51)$$

$$F_{dm} = \omega_b \Psi_{dm} = X_m (i_{ds} + i_{dr}) \quad (52)$$

Where  $X_{ls} = \omega_b L_{ls}$ ,  $X_{lr} = \omega_b L_{lr}$  and  $X_m = \omega_b L_m$  or in other words

$$F_{qs} = X_{ls} i_{qs} + F_{qm} \quad (53)$$

$$F_{qr} = X_{lr} i_{qr} + F_{qm} \quad (54)$$

$$F_{ds} = X_{ls} i_{ds} + F_{dm} \quad (55)$$

$$F_{dr} = X_{lr} i_{dr} + F_{dm} \quad (56)$$

From Equations (53)–(56), the currents can be expressed in terms of the flux linkages as:

$$i_{qs} = \frac{F_{qs} - F_{qm}}{X_{ls}} \quad (57)$$

$$i_{qr} = \frac{F_{qr} - F_{qm}}{X_{lr}} \quad (58)$$

$$i_{ds} = \frac{F_{ds} - F_{dm}}{X_{ls}} \quad (59)$$

$$i_{dr} = \frac{F_{dr} - F_{dm}}{X_{lr}} \quad (60)$$

Substituting Equations (57)-(60) into the voltage Equations (43)-(46),

$$V_{qs} = \frac{R_s}{X_{ls}} (F_{qs} - F_{qm}) + \frac{1}{\omega_b} \frac{dF_{qs}}{dt} + \frac{\omega_e}{\omega_b} F_{ds} \quad (61)$$

$$V_{ds} = \frac{R_s}{X_{ls}} (F_{ds} - F_{dm}) + \frac{1}{\omega_b} \frac{dF_{ds}}{dt} - \frac{\omega_e}{\omega_b} F_{qs} \quad (62)$$

$$0 = \frac{R_r}{X_{lr}} (F_{qr} - F_{qm}) + \frac{1}{\omega_b} \frac{dF_{qr}}{dt} + \frac{(\omega_e - \omega_r)}{\omega_b} F_{dr} \quad (63)$$

$$0 = \frac{R_r}{X_{lr}} (F_{dr} - F_{dm}) + \frac{1}{\omega_b} \frac{dF_{dr}}{dt} - \frac{(\omega_e - \omega_r)}{\omega_b} F_{qr} \quad (64)$$

Which can be expressed in state-space form as:

$$\frac{dF_{qs}}{dt} = \omega_b \left[ V_{qs} - \frac{\omega_e}{\omega_b} F_{ds} - \frac{R_s}{X_{ls}} (F_{qs} - F_{qm}) \right] \quad (65)$$

$$\frac{dF_{ds}}{dt} = \omega_b \left[ V_{ds} + \frac{\omega_e}{\omega_b} F_{qs} - \frac{R_s}{X_{ls}} (F_{ds} - F_{dm}) \right] \quad (66)$$

$$\frac{dF_{qr}}{dt} = -\omega_b \left[ \frac{(\omega_e - \omega_r)}{\omega_b} F_{dr} + \frac{R_r}{X_{lr}} (F_{qr} - F_{qm}) \right] \quad (67)$$

$$\frac{dF_{dr}}{dt} = -\omega_b \left[ -\frac{(\omega_e - \omega_r)}{\omega_b} F_{qr} + \frac{R_r}{X_{lr}} (F_{dr} - F_{dm}) \right] \quad (68)$$

And the torque equation can be written as:

$$T_e = \frac{3}{2} \left( \frac{P}{2} \right) \frac{1}{\omega_b} (F_{ds} i_{qs} - F_{qs} i_{ds}) \quad (69)$$

Equations (65)-(67), along with Equation (27), describe the complete model in state-space form where  $F_{qs}, F_{ds}, F_{qr}$  and  $F_{dr}$  are the state variables.

## VII. CONCLUSION

Induction motors controlled with Vector control method or Direct Torque Control methods give excellent dynamic performance. It gives decoupled control of torque and flux. But steady state equivalent circuit model can not be used for these methods. D-q models in various reference frame are used depending on the control requirements. Such dynamic models are very useful and also with slight modification can be used for speed estimation without speed sensor. Using these models enables us to estimate various parameters of induction motors just by measuring voltages and currents drawn by the motor.

## REFERENCES

- [1] A. E. Fitzgerald, et al., "Electric Machinery," 5th Ed., McGraw-Hill, 1990.
- [2] B. K. Bose, "Modern Power Electronics and AC Drives", Pearson publications, 2003.
- [3] G. R. Slemon, "Modeling Induction Machines for Electric Drives," IEEE Trans. on Industry Applications, Vol. 25, No. 6, pp. 1126-1131, Nov. 1989.
- [4] D. W. Novotney, et al.(Editor), "Introduction to Field Orientation and High Performance AC Drives," IEEE IAS Tutorial Course, 1986.
- [5] P.C. Krause, "Electric machines", prentice Hall, 1985.
- [6] R.Krishnan, "Electric Motor Drives Modeling, Analysis and Control", first edition, 2001Prentice-Hall International, Inc. Upper Saddle River, New Jersey 07458.
- [7] P. C. Sen, Principles of Electric Machines & Power Electronics, Wiley 1999
- [8] R.K.Rajput, "Electrical Machines," first edition, New York: McGraw- Hill, 1993, pp. 352-353

# Comparison of Exponential Companding Transform and CB-ACE Algorithm for PAPR Reduction in OFDM Signal

A. Neelam Dewangan , B. Mangal Singh,

**Abstract**—One of the main disadvantages of Orthogonal Frequency Division Multiplexing (OFDM) is its high peak-to-average power ratio (PAPR). As the simplest approach to reducing the PAPR, Clipping based Active Constellation Extension (CB-ACE) exhibits good practicability, and the repeated clipping-and-filtering (RCF) algorithm proposed by Jean Armstrong provides a good performance in PAPR reduction and out-of-band power's filtering. However, its way of filtering in frequency-domain requires RCF operations to control the peak regrowth, which degrades the bit error rate (BER) performance and greatly increases the computational complexity. Therefore, this paper put forward comparison of two existing techniques namely Exponential Companding Transform and CB-ACE Algorithm. The simulation results show that, exponential Companding Transform gives better result for PAPR Reduction and provides low complexity in Algorithm.

**Keywords**- CB-ACE, Exponential Companding Transform, OFDM, PAPR, RCF

## I. INTRODUCTION

As a promising technique, OFDM has been widely used in many new and emerging broadband communication systems, such as digital audio broadcasting (DAB), high-definition television (HDTV), wireless local area network (IEEE 802.11a and HIPERLAN/2). However, as the OFDM signals are the sum of signals with random amplitude and phase, they are likely to have large PAPR that requires a linear high-power- amplifier (HPA) with an extremely high dynamic range, which is expensive and inefficient. Furthermore, any amplifier nonlinearity causes intermodulation products resulting in unwanted out-of-band power.

A number of approaches have been proposed to deal with the PAPR problem, including clipping, clipping-and-filtering (CF), coding, companding transform, active constellation extension (ACE), selected mapping (SLM), partial transmit sequence (PTS), and so on [1]. Compared with other methods, clipping is the simplest and of good practicality. In particular, Jean Armstrong has proposed a RCF Algorithm which is also called Clipping Based Active Constellation Extension, which dramatically reduces the PAPR and limits the out-of band power to a low level , but excessively increases the computational complexity as well. Based on Jean Armstrong's method, this paper describes an improved approach which can provide good performance and lower complexity.

## II. DEFINITION OF OFDM SIGNALS AND PAPR

In OFDM, a block of N symbols,  $\{X_k, k=0, 1, \dots, N-1\}$ , is formed with each symbol modulating one of a set of subcarriers,  $\{f_n, n = 0, 1, \dots, N-1\}$  with equal frequency separation  $1/T$ , where T is the original symbol period. An inverse discrete Fourier transform (IDFT) can efficiently generate the multicarrier symbols. The IDFT of vector  $X[k] = [X_0, X_1, \dots, X_{N-1}]^T$  results in T/N spaced discrete time signal  $x[n] = [x_0, x_1, \dots, x_{N-1}]^T$ . Thus, the transmitted signal is

$$x_n = \frac{1}{\sqrt{N}} \int_{k=0}^{N-1} \exp j \frac{\omega \pi k n}{N} \quad 0 \leq k \leq N-1 \quad (1)$$

The PAPR of the transmitted signal can be written as

$$PAPR = \frac{\max_{0 \leq n \leq N-1} |x_n|^2}{E\{|x_n|^2\}} \quad (2)$$

The complementary cumulative distribution function (CCDF) is one of the most frequently used performance measures for PAPR reduction techniques, which denotes the probability that the PAPR of a data block exceeds a given threshold z. The CCDF of the PAPR of a data block of N symbols with Nyquist rate sampling is derived as

$$P(PAPR > z) = 1 - P(PAPR \leq z) = 1 - (1 - e^{-z})^N \quad (3)$$

## III. THE CB-ACE ALGORITHM

The basic principle of Clipping-Based Active Constellation Extension (CB-ACE) algorithm involves switching between the time domain and the frequency domain. Filtering and applying the ACE constraint in the frequency domain, after clipping in the time domain, both require iterative processing to suppress the subsequent regrowth of the peak power [3]

The CB-ACE algorithm is first used to clip the peak amplitude of the original Orthogonal Frequency Division Multiplexing (OFDM) signal. The clipping sample obtained after clipping the peak signals, denoted by  $C_n^{(i)}$ , is given by

$$C_n^{(i)} = \begin{cases} (|x_n^{(i)}| - A)e^{j\theta_n}, & |x_n^{(i)}| > A \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where  $C_n^{(i)}$  is the Clipping sample of ith iteration,  $x_n^{(i)}$  is the

oversampled OFDM signal, A is predetermined clipping level

The equation (4) says that the clipping sample is reduced to a value equal to zero when the peak amplitude of the original OFDM signal is less than or equal to the predetermined clipping level, A. If the peak amplitude of the original OFDM signal is greater than the predetermined clipping level, then the clipping sample is given by  $(|x_n^{(i)}| - A) e^{j\beta_n}$ , where the predetermined clipping level is subtracted from the oversampled OFDM signal and is then multiplied by an exponential value [3].

The predetermined clipping level, denoted by A, is related to the target clipping ratio,  $\gamma$  and is given by the equation 5 [3].

$$\gamma = \frac{A^2}{E[|x_n|^2]} \tag{5}$$

Where,  $\gamma$  is the target clipping ratio and A is predetermined clipping level. The clipping of the peak signal results to distortion of the original OFDM signal, namely In-Band Distortion and Out-of-Band Distortion . [3], [4]. The in-band distortion results in the system performance degradation and cannot be reduced, while, the out-of-band distortion can be minimized by filtering the clipped signals. The signal obtained after filtering the clipped signal is given by [3].

$$x^{(i+1)} = x^i + \mu \tilde{c}^{(i)} \tag{6}$$

where,  $\mu$  is positive real number ( $\mu$  varies from 0.1 to 1) and  $\tilde{c}^{(i)}$  is the anti-peak signal at the  $i^{\text{th}}$  iteration given by

$$\tilde{c}^{(i)} = T^{(i)} c^{(i)} \tag{7}$$

where,  $T^{(i)}$  is transfer matrix at the  $i^{\text{th}}$  iteration which is given by

$$T^{(i)} = \tilde{Q}^{*(i)} \tilde{Q}^{(i)} \tag{8}$$

where,  $\tilde{Q}^{*(i)}$  is conjugate of constellation order and  $\tilde{Q}^{(i)}$  is the constellation order

Though, the process of filtering completely eliminates the distortions caused by the clipping process, it introduces peak regrowth at some of the peak signals of the OFDM signal. The peak regrowth can be reduced by repeating the filtering process, which may again introduce some distortions. Therefore, the clipping and filtering processes are to be repeated until the peak signals are completely reduced. Hence, the Clipping-Based Active Constellation Extension (CB-ACE) Algorithm is also named as the Repeated Clipping and Filtering (RCF) process [3]

#### IV. THE EXPONENTIAL COMPANDING TRANSFORM ALGORITHM

The Exponential Companding Transform is also named as the Nonlinear Companding Transform. The idea of companding comes, from the use of companding in Speech Processing. Since, the Orthogonal Frequency Division Multiplexing (OFDM) signal is similar to that of the speech signal, in the sense that large signals occur very infrequently, the same companding technique can be used to improve the OFDM transmission performance. The key idea of the Exponential Companding Transform is to effectively reduce the Peak-to-Average Power Ratio (PAPR) of the transmitted or the companded Orthogonal Frequency Division Multiplexing

(OFDM) signals by transforming the statistics of the amplitudes of these signals into uniform distribution. The uniform distribution of the signals can be obtained by compressing the peak signals and expanding the small signals. The process of companding enlarges the amplitudes of the small signals, while the peaks remain unchanged. Therefore, the average power is increased and thus the Peak-to-Average Power Ratio (PAPR) can be reduced .

The Exponential Companding Transform can also eliminate the Out-of-Band Interference (OBI), which is a type of distortion caused by clipping the original OFDM signals. The other advantage of the companding transform is that, it can maintain a constant average power level. The proposed scheme can reduce the PAPR for different modulation formats and sub-carrier sizes without increasing the system complexity and signal bandwidth. The Exponential Companding Transform also causes less spectrum side-lobes.

The original Orthogonal Frequency Division Multiplexing (OFDM) signal is converted into the companded signal by using the Exponential Companding Transform. The companded signal obtained by using the Exponential or Nonlinear Companding Transform is given by the equation 9

$$H(x) = \text{sgn}(x) \sqrt[2]{\alpha \left[ 1 - \exp\left(-\frac{x^2}{\sigma^2}\right) \right]} \tag{9}$$

Where,  $h(x)$  – Companded Signal obtained by Exponential Companding Transform,  $\text{sgn}(x)$  -sign Function,  $\alpha$  – Average Power of Output Signals,  $x$ -original OFDM signal. The average power of the output signals, denoted by  $\alpha$ , is required in order to maintain the average amplitude of both the input and output signals at the same level. The average power of the output signals is given by the equation 10.

$$\alpha = \frac{E[|s_n|^2]}{E \left[ \sqrt[2]{\alpha \left[ 1 - \exp\left(-\frac{|s_n|^2}{\sigma^2}\right) \right]} \right]^2} \tag{10}$$

Where,  $\alpha$  – Average Power of Output Signals

$d$  – Power of the amplitude of the Companded Signal

#### V. SIMULATION RESULTS

The Peak-to-Average Power Ratio (PAPR) of the original Orthogonal Frequency Division Multiplexing (OFDM) signal i.e., the PAPR is to be calculated by using the equations (1), (2) and (3).

From the Figure 1, the Peak-to-Average Power Ratio (PAPR) of the original Orthogonal Frequency Division Multiplexing (OFDM) signal is equal to 11.8 dB with a Complimentary Cumulative Distribution Function (CCDF) of  $10^{-2}$  or 0.01. The Peak-to-Average Power Ratio (PAPR) of the

original Orthogonal Frequency Division Multiplexing (OFDM) signal is very high, which is evident from the Screen Shot 2.1. The high PAPR results to the increase in the complexity of the Analog-to-Digital Convertors (ADCs) and Digital-to-Analog Convertors (DACs), also reduces the efficiency of the power amplifiers.

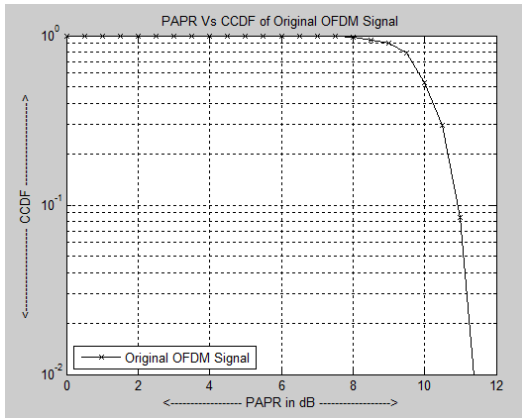


Figure1 – PAPR Vs CCDF of Original OFDM Signal

The Peak-to-Average Power Ratio (PAPR) by the Clipping-Based Active Constellation Extension (CB-ACE) algorithm is to be calculated for the Orthogonal Frequency Division Multiplexing (OFDM) signal which is obtained after filtering the clipped signal i.e., the PAPR is to be calculated for the equation (5) by using the equations (1), (2) and (3).

The Complimentary Cumulative Distribution Function (CCDF) by the Clipping-Based Active Constellation Extension (CB-ACE) algorithm is to be calculated for the Orthogonal Frequency Division Multiplexing (OFDM) signal which is obtained after filtering the clipped OFDM signal.

From the Figure 2, the Peak-to-Average Power Ratio (PAPR) of the Orthogonal Frequency Division Multiplexing (OFDM) signal obtained by using the Clipping-Based Active Constellation Extension (CB-ACE) algorithm is equal to 10 dB, 8.5 dB and 8.0 dB for the target clipping ratios of 0 dB, 2 dB and 4 dB respectively with a Complimentary Cumulative Distribution Function (CCDF) of  $10^{-2}$  or 0.01.

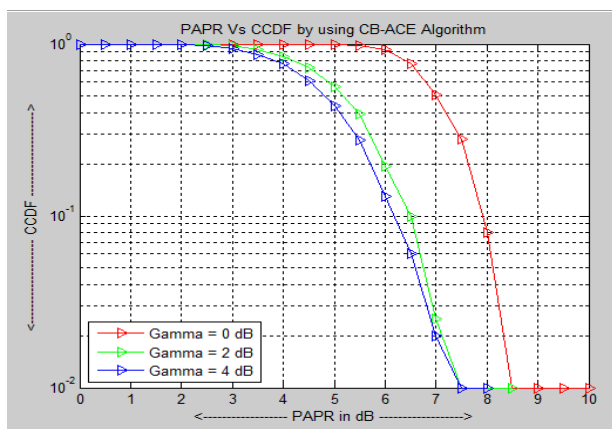


Figure 2 – PAPR Vs CCDF by using CB-ACE Algorithm  
(For Different Target Clipping Ratios)

The Peak-to-Average Power Ratios is increasing as the target clipping ratios is decreasing i.e., minimum PAPR cannot be achieved, when the target clipping level is set

below an initially unknown optimum value, which results to low clipping ratio problem.

The other problems faced by the Clipping-Based Active Constellation Extension (CB-ACE) algorithm are Out-of-Band Interference (OBI) and peak regrowth. Here, the Out-of-Band Interference (OBI) is a form of noise or an unwanted signal, which is caused when the original Orthogonal Frequency Division Multiplexing (OFDM) signal is clipped for reducing the peak signals which are outside to the predetermined area and the peak regrowth is obtained after filtering the clipped signal. The peak regrowth results to, increase in the computational time and computational complexity.

From Figure 3, the original Orthogonal Frequency Division Multiplexing (OFDM) signal is companded i.e., the peak signals of the OFDM signal are compressed and the small signals of the OFDM signal are expanded by using the Exponential Companding Transform for different powers of the amplitude of the companded signals i.e., for  $d = 1$ ,  $d = 2$  and  $d = 3$ .

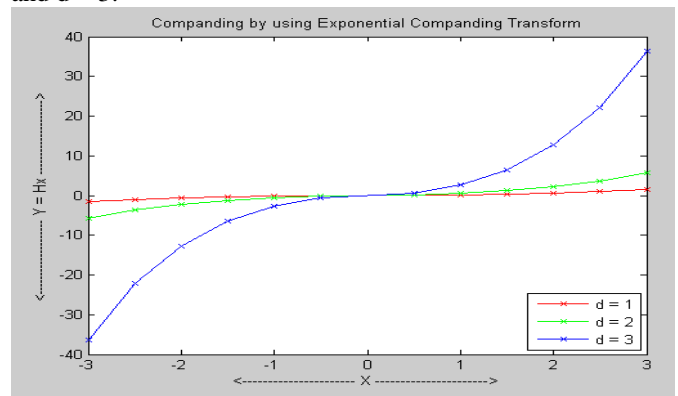


Figure 3 – Companding of OFDM Signal by using Exponential Companding Transform

The Peak-to-Average Power Ratio by using the Exponential Companding Transform is to be calculated for the Orthogonal Frequency Division Multiplexing (OFDM) signal which is obtained after compressing the peak signals and expanding the small signals i.e., PAPR is to be calculated for the equation (9) by using the equations (1), (2) and (3).

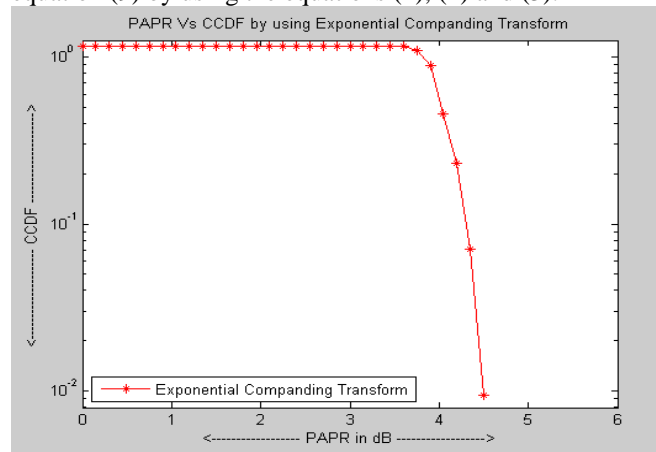


Figure 4 – PAPR Vs CCDF by using Exponential Companding Transform

From the Screen Shot 3.4, the Peak-to-Average Power Ratio (PAPR) of the Orthogonal Frequency Division Multiplexing

(OFDM) signals obtained by using the Exponential Companding Transform is reduced to 4.5 dB with a Complimentary Cumulative Distribution Function (CCDF) of  $10^{-2}$  or 0.01

Table 5.1 – Comparison of PAPR (in dB) and CCDF for different techniques

Different Techniques	PAPR (in dB)	CCDF
Original OFDM Signal	11.8	$10^{-2}$ or 0.01
Clipping-Based Active Constellation Extension (CB-ACE) Algorithm	10.0 (For $\gamma = 0$ dB) 8.5 (For $\gamma = 2$ dB) 8.0 (For $\gamma = 4$ dB)	$10^{-2}$ or 0.01
Exponential Companding Transform	4.5	$10^{-2}$ or 0.01

From the table 5.1, the Peak-to-Average Power Ratio of the Orthogonal Frequency Division Multiplexing systems is reduced or minimized by using the existing methods namely Clipping-Based Active Constellation Extension (CB-ACE) and the Exponential Companding Transform Algorithm at a Complimentary Cumulative Distribution Function of  $10^{-2}$  or 0.01.

## VI. CONCLUSIONS

The Clipping-Based Active Constellation Extension (CB-ACE) Algorithm reduces the high Peak-to-Average Power Ratio (PAPR) by clipping and filtering the original OFDM signal. The CB-ACE Algorithm results to peak regrowth, Out-of-Band Interference (OBI), low clipping ratio problem, increase in the Bit Error Rate (BER) and decrease in the Signal-to-Noise Ratio (SNR).

The Exponential Companding Transform improves the Bit Error Rate (BER) and minimizes the Out-of-Band Interference (OBI) in the process of reducing the Peak-to-Average Power Ratio (PAPR) effectively by compressing the peak signals and expanding the small signals. The improved BER transmits the data via a transmission channel with fewer errors, while the minimized OBI reduces the effects caused by clipping.

Hence, by reducing the Peak-to-Average Power Ratio (PAPR), the complexity of the Analog-to-Digital Converter (ADC) and Digital-to-Analog Converter (DAC) can be reduced. The reduced Peak-to-Average Power Ratio (PAPR) also increases the efficiency of the Power Amplifiers.

## REFERENCES

- [1] Seung Hee Han, Jae Hong Lee. An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. *Wireless Communications, IEEE*, Vol.12, Issue 2, pp.56–65, April, 2005
- [2] J. Armstrong. New OFDM Peak-to-Average Power Reduction Scheme. *IEEE cnf vehicular technology*, Vol.1, pp.756-760, May,2001
- [3] J. Armstrong. Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering. *Electronics letters*, Vol.38, No.5, pp.246-247, Feb.2002
- [4] IEEE Standard, 802.11a. Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. *IEEE*, 1999

# Real Time Implementation Of Stereo Matching Algorithm Using Matlab

A. Khyati N. Patel, B. Mrs. Sameena Zafar

**Abstract**— To realize the accuracy and stability of the real time binocular image matching, this paper design a stereo matching algorithm based on window & horizontal line based method. Firstly, apply median and homomorphic filtering for real time stereo image preprocessing to minimize noise and enhance image contrast. Secondly, apply the corner feature extraction and matching method based on window & horizontal line based method for obtaining the characteristic information of real time target. And then apply the constraint condition to eliminate the error matched points, so as to improve the real time image matching stability. Finally, the experiments verify the algorithm proposed in this paper, and the experimental results have shown the effectiveness of this algorithm.

**Index Terms**—Image Processing, Stereo matching, Window based method, horizontally line based method.

## I. INTRODUCTION

In recent years, it is an important research direction applying visual image technology for underwater target and environment detection. With the development of the theory of binocular vision, underwater binocular vision technology has been used in many fields, such as marine resources exploration, underwater target detection, marine ecological environment

Binocular vision, which is inspired by human visual process, computes the disparity between correspondence points in images captured by two cameras for distance measurement, and then recovers the depth information of the object. Considering the features of underwater environment, basis for applying binocular vision technology can better perceive underwater environment information. It provides the theory robot understanding the underwater environment and realizing the navigation and positioning. The image matching is one of the key technologies to realize underwater binocular vision. And the result of the matching would affect directly the precision of object recognition and 3D scene reconstruction.

Image matching is a process of seeking the corresponding feature points in two different images which are in the same scene. Currently, there are many research results of underwater image matching. stereo matching methods are divided into region-based matching, phase-based matching and feature-based matching. It can obtain a dense disparity map by region-based matching method. But it is rarely used because of its large amount of computation. The parallax image obtained by phase based matching method can reach sub-pixel accuracy. However, the method is more sensitive to the distortion, and it is difficult to choose the size of the matching

window precisely. Feature-based matching method can greatly reduce the amount of the matching process calculation. Since it is more sensitive to the position changes, the precision of matching is higher. But there exists miss detection phenomenon. In the three matching methods, the more classical algorithms are Absolute Balance Search algorithm, normalized cross-correlation matching algorithms, image moments matching algorithm, the matching algorithms based on Harris corner points and so on.

Binocular vision is the process of recovering depth from two images with the same height, the same direction and a certain distance, similar to human vision principle. During the process, stereo matching is the key point, which means to find the correspondence pixels of the same physical spatial point on both images. Binocular stereo matching algorithm research falls into two categories. One is based on sparse points, and the other is based on dense points. The latter one is more accurate on image matching. There are many binocular dense-point matching methods in which the representatives are Birch Field algorithm and Yoon algorithm. Birch Field algorithm is used to match two gray-scale images. It applies dynamic programming into matching epipolar lines. It is efficient but inaccurate. Subsequent improvement method of Birch Field is based on each pixel's eight neighborhoods and improves the matching accuracy to some extent. In the study of color image matching, Yoon's method of adaptive window algorithm is a milestone.

Obtaining reliable depth maps, indicating distance of surface from the stereo camera pair, have importance in robotic applications and autonomous systems. Intelligent systems, which can move around by itself, could be developed by obtaining dept information from the sensors. Stereovision is the one of methods that can yield dept information of the scene. It uses stereo image pairs from two cameras to produce disparity maps that can be easily turn into dept maps.

## II. BRIEF REVIEW

Stereo matching is another technique that is well known for measure and it is very easy to understand and program. With stereo matching we can get the exact position of a target by two stereo pictures. However the traditional stereo matching needs a strict condition about camera, such as the axis of stereo camera (a pair of camera with the same properties) must be kept parallel and the height of cameras will be the same. These conditions limited the application of stereo matching. The position of a point  $P$  in 3D space can be measured by traditional stereo matching that computes the coordinate of  $P$ ,  $(x, y, z)$ , from two pictures called left image and right image taken by two cameras separately (Figure 1). Based on the



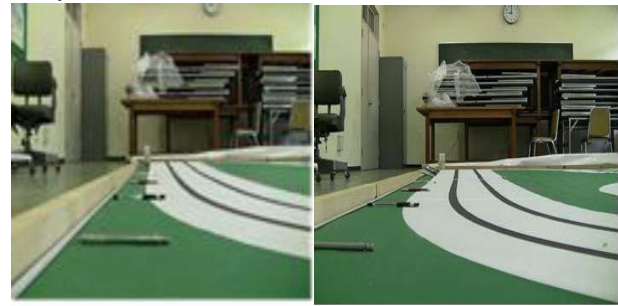
triangulation from two views the axis of the camera lenses must be paralleled to axis z, and the two cameras must be kept the same height, that is, the same y.

Figure:2Coordinates in xy-plane with point P

$$y = \frac{m_l}{f_l} \cdot \frac{y_l \cdot d}{\frac{m_l \cdot x_l}{f_l} - \frac{m_r}{f_r}} \tag{5}$$

$$z = \frac{d}{\frac{m_l \cdot x_l}{f_l} - \frac{m_r}{f_r}} \tag{6}$$

Figure 3(a) and (b) show the left image and the right image taken by two different cameras



(a) The image from left Camera (b) The image from right Camera

Figure 3. Stereo images for computation (targets are the pens on the table)

Figure:1Left and Right images with projection point P

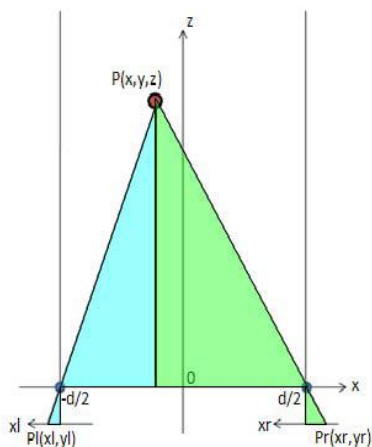
In Figure 1 and 2 the point P is our target we want to know  $P(x, y, z)$  exactly. For convenience we only show the x-z plane in Figure 2. Assume that the two cameras have the same focus distance  $f$ , the distance between the two cameras is  $d$ ,  $pl(x_l, y_l)$  denotes the coordinates of P in the left image and  $pr(x_r, y_r)$  denotes the coordinate of P in right image, we can get  $P(x, y, z)$  by,

$$x = \frac{d}{2} \cdot \frac{x_l + x_r}{x_l - x_r} \tag{1}$$

$$y = \frac{y_l \cdot d}{x_l - x_r} \tag{2}$$

$$z = \frac{f \cdot d}{m(x_l - x_r)} \tag{3}$$

$$x = \frac{d}{2} \cdot \frac{\frac{m_l \cdot x_l}{f_l} + \frac{m_r \cdot x_r}{f_r}}{\frac{m_l \cdot x_l}{f_l} - \frac{m_r \cdot x_r}{f_r}} \tag{4}$$



III. PROPOSED METHODS

Region Based Stereo Algorithms:

a) Global Error Energy Minimization by Smoothing Functions

In this method, we used block-matching technique in order to construct an Error Energy matrix for every disparity. Lets denote left image in RGB format by  $L(i, j, c)$ , denote right image in RGB format by  $R(i, j, c)$  and error energy by  $e(i, j, d)$ . For  $n \times m$  window size of block matching, error energy  $e(i, j, d)$  can be expressed by,

$$e(i, j, d) = \frac{1}{3 \times n \times m} \sum_{x=i}^{i+n} \sum_{y=j}^{j+m} \sum_{k=1}^3 (L(x, y + d, k) - R(x, y, k))^2 \tag{7}$$

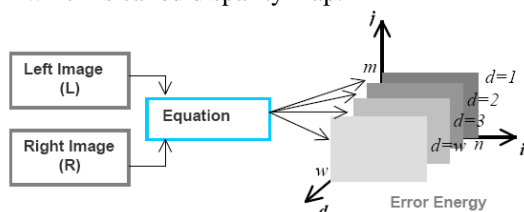
where,  $c$  represents RGB components of images and takes value of  $\{1,2,3\}$  corresponding to red, blue and green.  $d$  is the disparity. For a predetermined disparity search range ( $w$ ), every  $e(i, j, d)$  matrix respect to disparity is smoothed by applying averaging filter many times. Averaging filter (linear filter) removes very sharp change in energy which possibly belongs to incorrect matching. An other important properties of repeating application of averaging filter is that it makes apparent global trends in energy. (Local filtering in iterations could solve a global total variation optimization problem) Considering global trend in error energy naturally makes this algorithm a region based algorithm. For  $n \times m$  window size, averaging filtering of  $e(i, j, d)$  can be expressed by following equation,



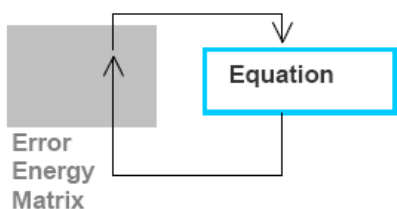
$$\tilde{e}(i, j, d) = \frac{1}{n+m} \sum_{x=i}^{i+n} \sum_{y=j}^{j+m} e(x, y, d) \quad (8)$$

After iterative application of averaging filtering to error energy for each disparity, we selected the disparity ( $d$ ), which has minimum error energy  $e\sim(i, j, d)$  as the most reliable disparity estimation for pixel ( $i, j$ ) of disparity map. Let's write basic steps of algorithm more properly,

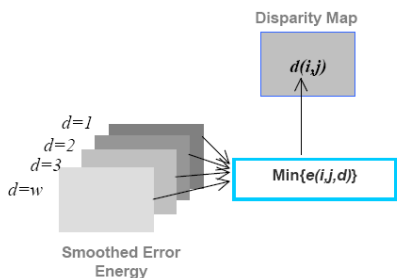
- Step 1:** For every disparity  $d$  in disparity search range, calculate error energy matrix.
- Step 2:** Apply average filtering iteratively to every error matrix calculated for a disparity value in the range of disparity search range.
- Step 3:** For every ( $i, j$ ) pixel, find the minimum error energy  $e\sim(i, j, d)$ , assign its disparity index ( $d$ ) to  $d(i, j)$  which is called disparity map.



(a) Construction error energy matrix



(b) Smoothing energy matrix for every disparity values



(c) Disparity map generation by minimum energy points

Figure 8. Method using global error energy minimization by smoothing functions

**b) Line Growing Based Stereo Matching**

Also proposed an algorithm based on region growing. In this manner, we consider region-growing mechanism in two phases operation. First phase, finding root point to grow region (*Root Selection process*) and the second phase, growing region for a root point corresponding to predefined rule. (*Region Growing process*) Our rule for associating a point to root point in the growing process is to have lower error energy than a predetermined threshold of error energy (*Line Growing Threshold*). In our application, being associated to a root points means to have the same disparity by root point. Thereby, the region emerged from all associated points have a disparity

value. Actually, we should call the algorithm as disparity growing. Let's generally express steps of the algorithm in a list,

- Step 1:** (*Root Selection process*) Select a point, which isn't belonging to any grown region and find its disparity using energy function equation (1). Set it root point and set its disparity to region disparity then go to step 2. If you didn't find any disparity with lower enough error energy, repeat this step for the next point.
- Step 2:** (*Region Growing process*) Calculates error energy of neighbor points just for root point disparity, which was called region disparity. If it is lower than the predetermined error energy threshold, associate this point to region.
- Step 3:** Proceed the Step 2 until region growing any more. In the case that region growing is completed, turn back to step 1 to find out new root point to repeat these steps. When all points in image processed, stop the algorithm. Grown disparity regions composes disparity map  $d(i, j)$ .

In order to reduce complexity of the algorithm, we allow the region growing in the direction of rows since disparity of stereo image is only in row directions. So, only one neighbor, which is the point after searched point, is inspected for region growing.

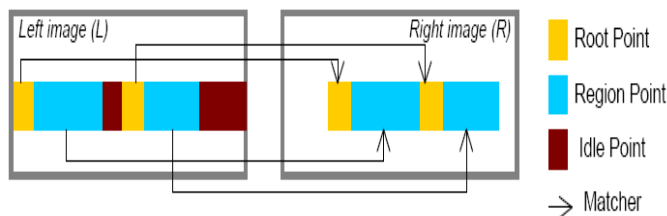
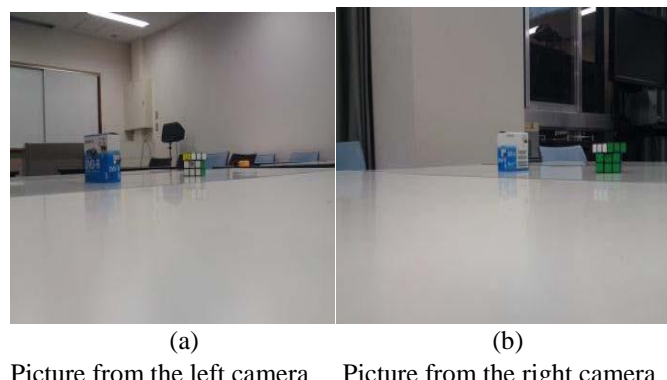


Figure: 9 Method using line growing

**IV. PROBLEM STATEMENT**

In our work a number of cameras are located around of target and the two pictures taken by each pair are used for stereo matching calculation. It is impossible to compute the coordinate of target because the axes of camera lens are probably not paralleled each other. Assume these cameras having the same height we proposed a stereo axes correction algorithm and new equations to solve the unparallelled lens axes problem.



Picture from the left camera      Picture from the right camera

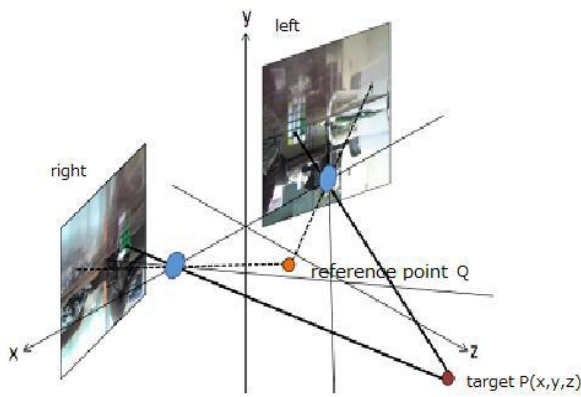


Figure 10. Stereo axes correction

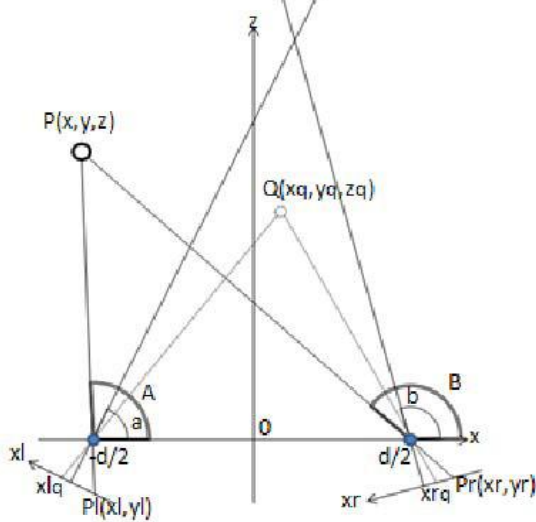


Figure 11. The principle of stereo axes correction  
The Rubik's cube both in the left and right images of Figure 10 is our measured target. We do know the two axes of camera lens are not paralleled but we are not sure exactly how much they incline each other. In order to find out the lens axes inclined angle  $a$  and  $b$  of Figure 11 we pick up a known point  $Q(x_q, y_q, z_q)$  as the reference point whose depth value  $z_q$  and the distance between camera and  $Q$  can be measured by any measure tool. Figure 12 and (9) gives the method to measure the point  $Q$ .

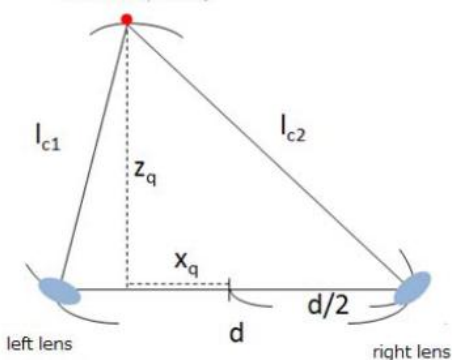


Figure 12. The Measure of the reference point

$$x_q = \frac{l^2 \cdot c_1 - l^2}{2 \cdot d} \tag{9}$$

$$z_q = \sqrt{l^2 \cdot c_1 - (x_q + \frac{d}{2})^2} \tag{10}$$

Now the known  $Q(x_q, y_q, z_q)$  into equations, where  $x_{lq}$  is coordinate  $x$  of  $Q$  in the left image and  $x_{rq}$  is coordinate  $x$  of  $Q$  in the right image,

$$a = \tan^{-1} \left( \frac{z_q}{x_q + \frac{d}{2}} \right) + \tan^{-1} \left( \frac{x_{lq} \cdot m}{f} \right) \tag{11}$$

$$b = \tan^{-1} \left( \frac{z_q}{x_q - \frac{d}{2}} \right) + \tan^{-1} \left( \frac{x_{rq} \cdot m}{f} \right) \tag{12}$$

$A$  and  $B$  that are inclined angles of target  $P$  in the left image and the right image can be expressed by

$$A = a - \tan^{-1} \left( \frac{x_{l \cdot l}}{f} \right) \tag{13}$$

$$B = a - \tan^{-1} \left( \frac{x_{r \cdot l}}{f} \right) \tag{14}$$

Therefore the coordinate of target  $P$  can be obtained by

$$X = -\frac{d}{2} \cdot \frac{\tan(A) + \tan(B)}{\tan(A) - \tan(B)} \tag{15}$$

$$y = \frac{y_l \cdot m \cdot \sqrt{(x + \frac{d}{2})^2}}{\sqrt{f^2 + (m \cdot x_l)^2}} \tag{16}$$

$$z = \tan(A) \cdot (x + \dots) \tag{17}$$

If the properties of the left camera and the right camera are different, let  $f_l, m_l$  and  $f_r, m_r$  stand for the focus distance and coefficients  $m$  of the left camera and the right camera separately,  $a, b, A$  and  $B$  are rewritten as follows,

$$a = \tan^{-1} \left( \frac{z_q}{x_q + \frac{d}{2}} \right) + \tan^{-1} \left( \frac{x_{lq} \cdot m}{f_l} \right) \tag{18}$$

$$b = \tan^{-1} \left( \frac{z_q}{x_q - \frac{d}{2}} \right) + \tan^{-1} \left( \frac{x_{rq} \cdot m}{f_r} \right) \tag{19}$$

$$A = a - \tan^{-1} \left( \frac{x_{l \cdot n}}{f_l} \right) \tag{20}$$

$$B = b - \tan^{-1} \left( \frac{x_{r \cdot n}}{f_r} \right) \tag{21}$$

Having the new equations not worry about whether the lens axes of cameras are paralleled or not anymore. And the shadow area of figure become more widely than before, if necessary we can incline leans axes intentionally to get better broad views. Meanwhile in the new method the reference point plays an important role to correct the unparalleled lens axes.

A critical issue in stereo matching is to measure the similarity (dissimilarity) between correspondences, which is calculated as a *matching cost*. Common matching costs defined based on the brightness constancy assumption, i.e., scene points have similar intensities in different views, are Absolute Difference and Squared Difference. Using the matching cost, many local and global stereo methods have been proposed to improve the matching accuracy, substantially.

Other matching cost functions obtain robustness to radiometric differences by removing or relaxing the brightness constancy assumption. Hirschmuller evaluated many of them such as Normalized Cross-Correlation (NCC), rank and census transforms, LoG and mean filters. As a more complicated measure, mutual information method can handle more complex radiometric transformations. When globally

reasoning the image radiometric transformation, mutual information method is comparably sensitive to local variations such as vignetting.

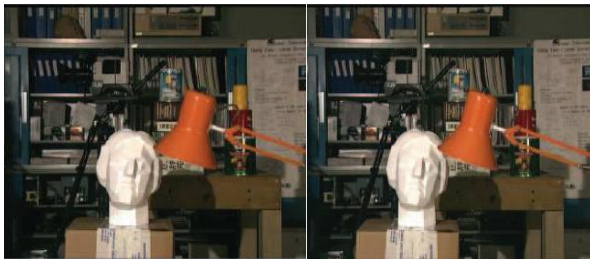


Figure:13. Matching the Tsukuba (a) left image and (b) its right image with a global intensity bias

V. WORK

The major innovative point is to combine color aggregation with local disparity estimation and adaptive window matching. It is able to accomplish a better matching accuracy while effectively reducing the time complexity thus improves the performance of the algorithm.

Compared with land image, there are more significant differences in underwater image quality. Medium's strong absorption of light and scattering properties, the underwater image has the characteristics of low contrast, high ambiguity and low image pixel resolution.

In underwater image based on stereo matching algorithm is already present and here we can see that which output will come if underwater image is based on window-based adaptive correspondence search algorithm but we can use horizontal line based method.

Using window and Horizontal line we use the following steps: Take the disparity map estimated from the above step as the initial value, now we use window-based correspondence method to optimize the result.

For each pixel  $p$  and its neighborhood  $N_p$  in reference image, the corresponding pixel  $p_d$  and its neighborhood  $N_p$  in target image, define the dissimilarity  $E(p, p_d)$  between the two windows:

$$E(p, p_d) = \frac{\sum_{q \in N_p, q_d \in N_{p_d}} w(p, q)w(p_d, q_d)e(q, q_d)}{\sum_{q \in N_p, q_d \in N_{p_d}} w(p, q)w(p_d, q_d)} \quad (22)$$

Where  $e(q, q_d)$  is absolute difference and  $w(p, q)$  is the adaptive weight:

$$e(q, q_d) = \min\{\sum_{c \in \{r, g, b\}} |I_c(q) - I_c(q_d)|, T\} \quad (23)$$

$$w(p, q) = f(\Delta c_{pq}, \Delta g_{pq})$$

$\Delta c_{pq}$  and  $\Delta g_{pq}$  are color similarity and geometric proximity.

To better understand depth and disparity relation, let see stereo projection representation illustrated in the Figure 14. By considering the figure, one can derive relation between dept ( $Z$ ) and disparity ( $d$ ) by using basic geometrical calculations as following.

$$Z(i, j) = f \cdot \frac{T}{d(i, j)} \quad (24)$$

If real location of object surface projected at pixel ( $i, j$ ) is willing to calculate, following formulas can be used in calculation of ( $X, Y$ ) points after calculation of the  $Z$ .

$$X = \frac{(Z-f)}{f} \cdot i \quad \text{and} \quad Y = \frac{(Z-f)}{f} \cdot j \quad (25)$$

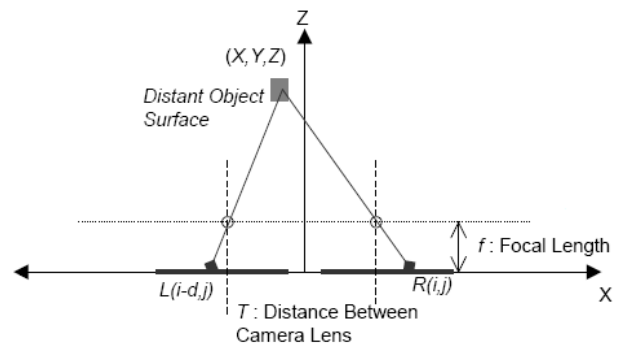


Figure 14. Representation of the stereo projection

In order to obtain smoother depth map to be used in applications such as robot navigation and recent trend for vision in various engineering application.

For that we take a various left and right angle with putting different distance of the camera/webcam and determine the disparity map for image, if time permits then try for real time interface with camera/webcam to computer system and prepare the stereo matching of the images and determine using smoothing with and without consider the reliability of non estimating pixel of images, also determine RMS error, time to be taken for execution in our system.

Our work progress chart show in figure 15.

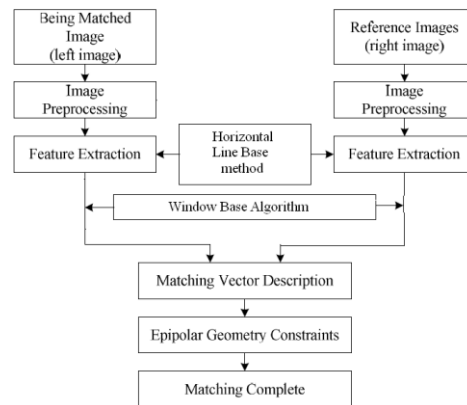


Figure:15 work progress chart



Left camera Image right camera image Disparity Map

Figure:16 Window Based Method image and Disparity Map



Left camera image      right camera image      Disparity Map  
 Figure:17 Horizontal Line Based Method image and Disparity Map

## VI. CONCLUSION

According to the stability demand of the real time binocular image processing, this paper introduces an underwater image matching method based horizontal line base method. Because there is special corresponding relation between binocular images, we apply epipolar geometry constrain to eliminate the error matched points, enhancing the matching efficiency. In the study, we find that, in this special real time environment, the preprocessing quality of the image and the scale sensitivity of the algorithm has a great influence on the image matching accuracy. So how to further improve the image preprocessing effect and reduce the algorithm scale sensitivity is our future research.

HL	high light and low noise	Output	low light and high noise	Output
0.6	75	98.4%	200	97.2%
0.7	80	90.5%	225	89.1%

## REFERENCES

- [1] Huibin Wang, Hongye Sun, Jie Shen, Zhe Chen, "A Research on Stereo Matching Algorithm for Underwater Image", 2011 4th International Congress on Image and Signal Processing, 978-1-4244-9306-7/11/\$26.00 ©2011 IEEE, pp.850-854.
- [2] Shujun Zhang, Jianbo Zhang, Yun Liu, "A Window-Based Adaptive Correspondence Search Algorithm Using Mean Shift and Disparity Estimation", 2011 International Conference on Virtual Reality and Visualization, DOI 10.1109/ICVRV.2011.47, 978-0-7695-4602-5/11 \$26.00 © 2011 IEEE, pp.319-322.
- [3] Ningping Sun, Hiroshi Nagaoka, Takuya Shigemoto, "A Correction Algorithm for Stereo Matching with Web Cameras", 2011 International Conference on Broadband and Wireless Computing, Communication and Applications, DOI 10.1109/BWCCA.2011.88, 978-0-7695-4532-5/11 \$26.00 © 2011 IEEE, pp.543-548.
- [4] Jian Wei, Shigang Wang, Liwei Chen, Tianxiao Guan, "Adaptive Stereo Video Object Segmentation Based on Depth and Spatio-Temporal Information", 2009 World Congress on Computer Science and Information Engineering, DOI 10.1109/CSIE.2009.422, 978-0-7695-3507-4/08 \$25.00 © 2008 IEEE, pp. 140-144
- [5] M. Kaaniche, W. Miled, B. Pesquet-Popescu, A. Benazza-Benyahia and J.- C. Pesquet, "Dense Disparity Map Representations For Stereo Image Coding", ICIP 2009, 978-1-4244-5654-3/09/\$26.00 ©2009 IEEE, pp.725-728.
- [6] Baris Baykant ALAGÖZ, "Obtaining Depth Maps From Color Images By Region Based Stereo Matching Algorithms", *OncuBilim Algorithm And Systems Labs. Vol.08, Art.No:04,(2008),pp.1-13.*
- [7] Andrea Fusiello, Emanuele Trucco, Alessandro Verri, "A compact algorithm for rectification of stereo pairs", *Machine Vision and Applications (2000) Springer-Verlag 2000*, pp.16–22.
- [8] Ke Zhang, Jiangbo Lu, Gauthier Laffruit1, Rudy Lauwereins, Luc Van Gool, "Robust Stereo Matching With Fast Normalized Cross-Correlation Over Shape-Adaptive Regions", ICIP 2009, 978-1-4244-5654-3/09/\$26.00 ©2009 IEEE, pp.2357-2360.
- [9] N. Grammalidis and M.G. Strintzis, "Disparity and Occlusion Estimation in Multicocular Systems and their Coding for the Communication of Multiview Image Sequences," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 328–344, June 1998.

- [10] N. Grammalidis and M.G. Strintzis, "Disparity and Occlusion Estimation in Multicocular Systems and their Coding for the Communication of Multiview Image Sequences," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 328–344, June 1998.
- [11] Geoffrey McLachlan and Thriyambakam Krishnan. *The EM Algorithm and Extensions*. John Wiley & Sons, New York, 1996.
- [12] L. Hong and G. Chen. Segment-based stereo matching using graph cuts. In *CVPR*, pages I: 74–81, 2004.
- [13] M. Bleyer and M. Gelautz. A layered stereo matching algorithm using image segmentation and global visibility constraints. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3):128–150, May 2005
- [14] J. Sun, N. Zheng, and H. Shum. Stereo matching using belief propagation. In *ECCV*, page II: 510 ff., 2002.
- [15] Q. Yang, L. Wang, R. Yang, H. Stew'enius, and D. Nist'er. Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling. Accepted to *CVPR* 2006.

# Digital FIR Band Pass Filter Implementation Using Remez Exchange Algorithm

A. Nirav R. Patel

**Abstract**—A filter may be required to have a given frequency response, or a specific response to an impulse, step, or ramp, or simulate an analog system. Depending on the response of the system, digital filters can be classified into Finite Impulse Response (FIR) filters & Infinite Impulse Response (IIR) filters. FIR Filters can be designed using frequency sampling or windowing methods. But these methods have a problem in precise control of the critical frequencies. In the optimal design method, the weighted approximation error between the actual frequency response and the desired filter response is spread across the pass-band and the stop-band and the maximum error is minimized, resulting in the pass-band and the stop-band having ripples. The peak error can be computed using a computer-aided iterative procedure, known as the Remez Exchange Algorithm.

**Index Terms**— FIR Filter, FDA tool, Taps, Remez Exchange Algorithm

## I. INTRODUCTION

FIR filters are a special kind of digital filters. They are non-recursive type of filter where the present output depends on the present input sample and the previous samples. The impulse response of FIR filter has finite number of non-zero terms [3]. Few characteristics of FIR which serve as their advantage are enlisted below [7]:

1. FIR filters have exactly linear phase.
2. FIR filters are always stable.
3. FIR filters may be realized in both recursive and non-recursive structures.

The structure of a general FIR filter is shown in fig 1. From this structure the transfer function of the filter can be easily described in z- domain as:

$$H(z) = b_0 + b_1z^{-1} + \dots \tag{1}$$

Where N= filter order.

FIR filters are used to design a filter where there is a requirement for a linear phase characteristic within the pass band of that filter. The performance of different methods depends on the similarities of its frequency response with the

desired frequency response. To follow this philosophy, there are three techniques to design FIR filters [4]. In this paper, the spectral responses of those methods are analyzed. Since this is

FIR system so it has an impulse response of finite duration length, N.

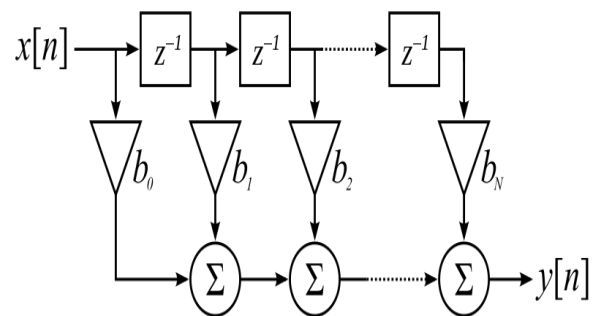


Fig. 1. Block diagram of FIR filter

The objective of most FIR coefficient calculation methods is to obtain values of such that the resulting filter meets the design specifications. Several methods are available for obtaining which are [1].

1. Windowing method
2. iterative method

Windowing method is much easier to implement. It is a straightforward approach. If we assume that  $H_d(e^{j\omega})$  to be ideal response and  $h_{fin}(n)$  to be corresponding infinite-duration impulse response sequence, then we can obtain a finite-duration causal impulse response by multiplying  $h_{fin}(n)$  with finite-duration “window”  $w(n)$

$$h_{fin}(n) \cdot w(n) = h(n) \tag{2}$$

Window functions for filter design should taper smoothly to a value of zero at each end. Two commonly used windows are the Hamming and the Blackman window. These windows are defined by the following equations:

$$w_{ham}(n) = 0.54 + 0.46 \cos\left(\frac{n\pi}{N}\right), -N \leq n \leq N \tag{3}$$

A. Works with the Electronics Engineering Department, Kalol Institute of Technology & Research center, Kalol, Gujarat, India. (E-mail: niravpatel34@gmail.com).

$$w_{black}(n) = 0.42 + 0.5 \cos\left(\frac{n\pi}{N}\right) + 0.08 \cos\left(\frac{2n\pi}{N}\right)$$

$$-N \leq n \leq N$$

(4)

The iterative method allows the design of optimal filters. Optimal filter means filter with constant equiripple in the pass band and the stop band. The most commonly applied algorithm used is the Remez Exchange Algorithm. This algorithm is being discussed below.

## II. REMEZ EXCHANGE ALGORITHM

A powerful technique which employs the Remez exchange algorithm to find the extreme frequencies has been developed. By knowing the locations of the extreme frequencies, it is a simple matter to work out the actual frequency response and the impulse response of filter. For given set of specifications the optimal method involves the following key steps: [2]

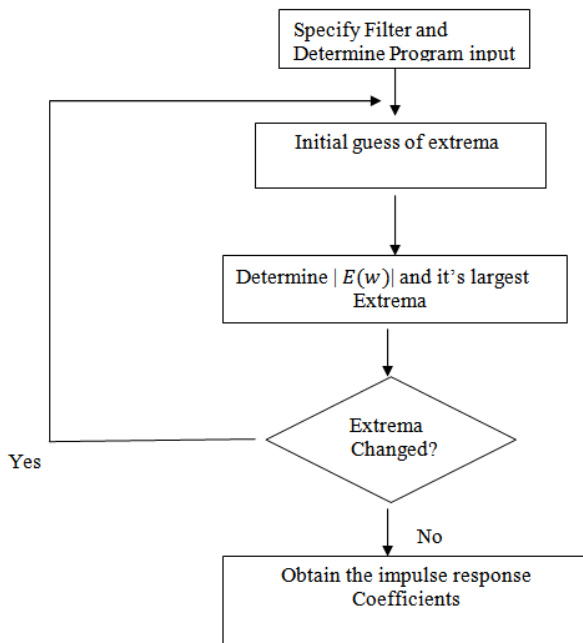


Fig. 2. Simplified flowchart of the Remez exchange algorithm

The heart of the Remez exchange algorithm is the first step where an iterative process is used to determine the extreme frequencies of a filter whose amplitude-frequency response satisfies the optimality condition. The result is called the polynomial of best approximation, Chebyshev approximation, or the minimax approximation. The Remez FIR Filter Design block implements the Parks-McClellan algorithm to design and apply a linear-phase filter with an arbitrary multiband magnitude response. The filter design, which uses the firpm function in Signal Processing Toolbox, minimizes the maximum error between the desired frequency response and the actual frequency response. Such filters are called equiripple due to the equiripple behavior of their

approximation error. The block applies the filter to a discrete-time input using the Direct-Form II Transpose Filter block. Parks-McClellan optimal FIR filter design Syntax

$$b = \text{firpm}(n, f, a)$$

(5)

$$b = \text{firpm}(n, f, a, 1)$$

(6)

Firpm designs a linear-phase FIR filter using the Parks-McClellan algorithm. The Parks-McClellan algorithm uses the Remez exchange algorithm and Chebyshev approximation theory to design filters with an optimal fit between the desired and actual frequency responses. The filters are optimal in the sense that the maximum error between the desired frequency response and the actual frequency response is minimized. Firpm exhibits discontinuities at the head and tail of its impulse response due to this equiripple nature.

$b = \text{firpm}(n, f, a)$ , returns row vector  $b$  containing the  $n+1$  coefficients of the order  $n$  FIR filter whose frequency-amplitude characteristics match those given by vectors  $f$  and  $a$ . The output filter coefficients (taps) in  $b$  obey the symmetry relation:

$$b(k) = b(n + 2 - k), \quad k = 1, \dots, n$$

(7)

## III. FILTER SPECIFICATION

The implementation of Remez FIR filter is done for the specifications characterizing an equiripple FIR design method for a Band pass response type. The following figure illustrates the magnitude frequency responses of a band pass filter, which passes a certain band of frequencies and attenuates lower and higher frequencies. In the figure, stop band edge frequency 1 indicates the maximum frequency of the lower frequency range that you want to attenuate, and stop band edge frequency 2 indicates the minimum frequency of the higher frequency range that you want to attenuate. The frequency range between pass band edge frequency 1 and 2 indicates the range of frequencies that can pass through the filter [4].

Filter performance parameter	Value
Stop band Frequency1 (Fstop1)	8000 Hz
Pass band Frequency1 (Fpass1)	9600 Hz
Pass band Frequency (Fpass2)	12800 Hz
Stop band Frequency2( Fstop2)	14400 Hz
Sampling frequency	48000 HZ

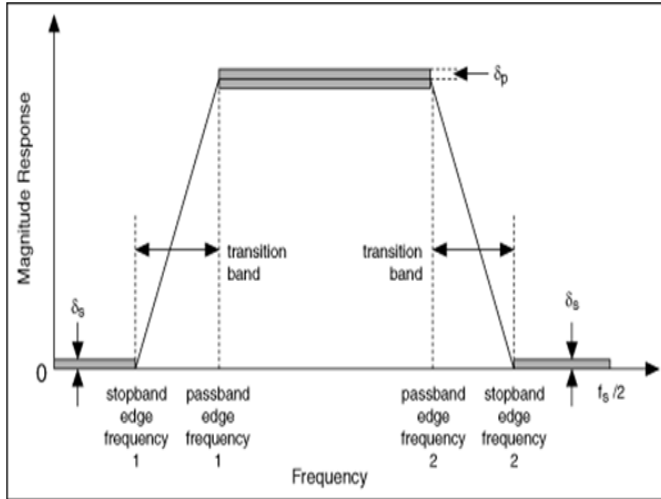


Fig. 3. Magnitude response specifications for a BPF

Notice the transition band between the passband and stopband frequencies. In an ideal design, a digital filter has a target gain in the passband and a zero gain ( $-\infty$  dB) in the stopband. In a real implementation, a finite transition region between the passband and the stopband, which is known as the transition band, always exists.  $\delta_p$  indicates the magnitude of the passband ripple, which equals the maximum deviation from the unity.  $\delta_s$  indicates the magnitude response of the stopband ripple, which equals the maximum deviation from zero. The gain of the filter in the transition band is unspecified. The gain usually changes gradually through the transition band from 1 (0 dB) in the passband to 0 ( $-\infty$  dB) in the stopband. The stop band ripple,  $A_s$  and the pass band ripple,  $A_p$ , in decibels are given as [7]

$$A_s \text{ (stop band attenuation)} = -20 \log_{10} \delta_s \quad (8)$$

$$A_p \text{ (pass band ripple)} = 20 \log_{10} (1 + \delta_p) \quad (9)$$

#### IV. FILTER IMPLEMENTATION

The Filter Design and Analysis Tool is a powerful Graphical user interface (GUI) for designing and analyzing filters quickly. Filter Design and Analysis Tool enables you to design digital FIR or IIR filters by setting filter specifications, by importing filters from your MATLAB workspace, or by adding, moving or deleting poles and zeros. Filter Design and Analysis Tool also provides tools for analyzing filters, such as magnitude and phase response and pole-zero plots. Filter Design and Analysis Tool seamlessly

integrates additional functionality from other Math Works products [8].

The frequency specifications are as follows:

TABLE I  
FIR FILTER SPECIFICATION

In this paper, Remez exchange algorithm is used to meet the specifications given in table 1 using FDA tool. The results of this FIR filter design technique is then compared for different orders.

The magnitude responses of band-pass Equiripple FIR filters of length  $N=16$ ,  $N=32$  and  $N=64$  are shown in the Fig.3, Fig.4 and Fig.5.

The number of FIR taps tells us the amount of memory needed, the number of calculations required and the amount of filtering that it can do. Basically, the more taps in the filter results in better stop band attenuation (less of the part we want filtered out), less rippling (less variation in the pass band) and steeper roll-off (a shorter transition between pass band and the stop band) [2]. These points are demonstrated clearly in Fig.6. The magnitude responses of band-pass FIR filters using Remez exchange algorithm of length  $N=16$ ,  $N=32$  and  $N=64$  are shown in the Fig.6.

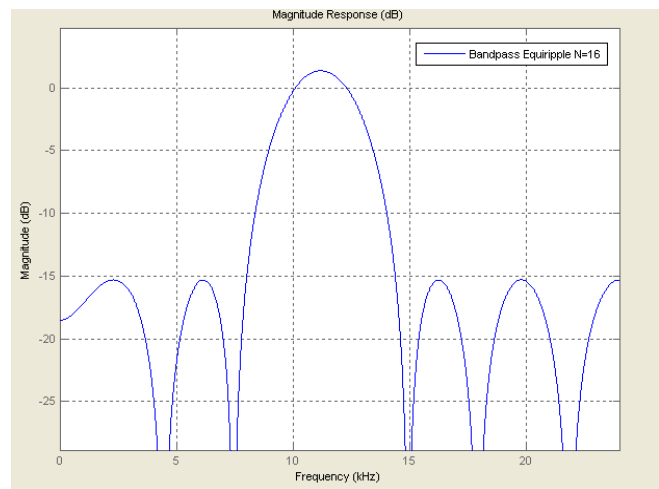


Fig. 3. Magnitude Response of Remez FIR filter of order 16

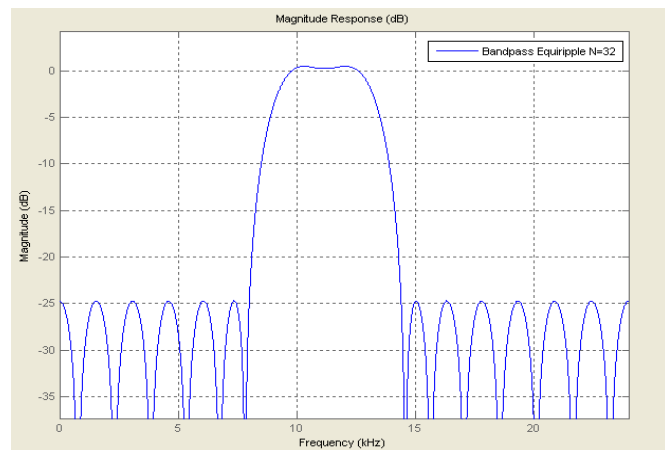


Fig. 4. Magnitude Response of Remez FIR filter of order 32

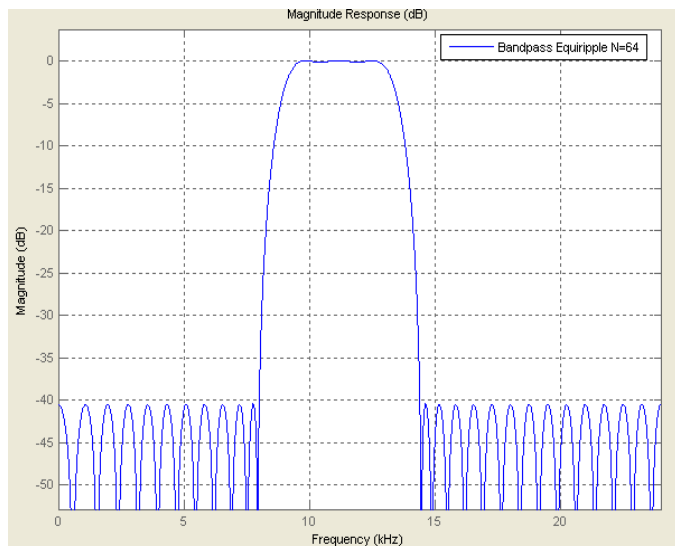


Fig. 5. Magnitude Response of Remez FIR filter of order 64

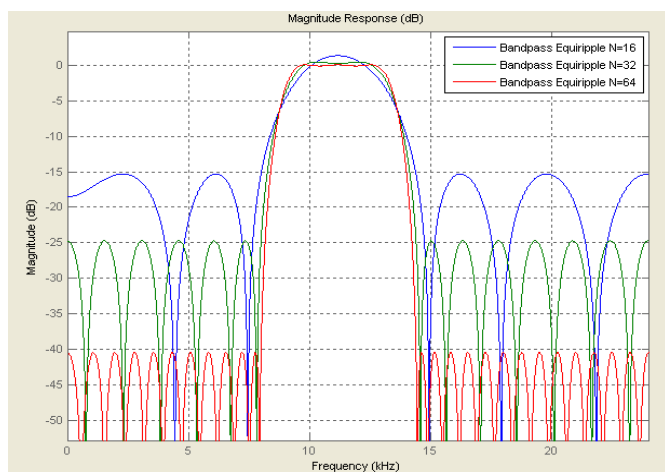


Fig. 6. Magnitude responses of band pass equiripple FIR filters of Length  $N=16$ ,  $N=32$  and  $N=64$

Remez exchange algorithm can be used to meet the specifications given in table 1 and their relative performance is shown in table 2 for different orders. While designing digital filter, order of filter (N) is always a matter of concern. In this paper, the performance of FIR filter design is measured in terms of order required to implement digital filter.

TABLE II  
PERFORMANCE COMPARISON OF FIR FILTER DESIGN FOR DIFFERENT ORDERS

Sr. No.	Order of Window(N)	Pass band Attenuation ( $\delta_p$ )	Stop band Attenuation ( $\delta_s$ )
1.	16	1.325 dB	15.532

2.	32	0.432 dB	25.325 dB
3.	64	0.267 dB	41.426 dB

### V. CONCLUSION

FIR filter using Remez Exchange Algorithm has been implemented on a MATLAB. The number of FIR taps tells us the amount of memory needed, the number of calculations required and the amount of filtering that it can do. Basically, the more taps in the filter results in better stop band attenuation (less of the part we want filtered out), less rippling (less variation in the pass band) and steeper roll-off (a shorter transition between pass band and the stop band). The FIR filter using Remez Exchange Algorithm is optimal in the sense that it minimizes the maximum error between the desired frequency response and actual frequency response. The use of Remez Exchange algorithm includes several advantages over other methods of calculating filter coefficients. This particular algorithm tremendously reduces the number of multipliers and adders being used in narrow-transition band linear-phase FIR filter. The overall process of synthesis is found to be very much faster than other methods known.

### REFERENCES

- [1] B.A.Shenoi, "Introduction to Digital Signal Processing & Filter design", A John Wiley & Sons, Inc., Publication, 2006, pp. 249-292.
- [2] Morgan Kaufmann, Salivahanan S., Vallavaraj A., Gnanapriya C.(2000), "Digital Signal Processing", Tata McGraw Hill.
- [3] N. G. Palan, "Digital Signal Processing", third edition, Tech-Max Publication, 2008.
- [4] Li Tan, "Digital Signal Processing-Fundamentals and Applications", 2008, Elsevier Inc., pp. 216-288.
- [5] J Forester W. Isen, "DSP for MATLAB™ and Lab VIEW™ Volume III: Digital Filter Design", 2008, pp.39-69.
- [6] S. M. Shamsul Alam and Md. Tariq Hasan, "Performance Analysis of FIR Filter Design by Using Optimal, Blackman Window and Frequency Sampling Methods", *International Journal of Electrical & Computer Sciences*, Vol 10, No -1, Dec 2009, pp. 13-18.
- [7] Animesh Panda, Shaileh Kumar Agrawal, T. Siva Kumar, T. Usha & Satish Kumar Baghmar, "FIR Filter Implementation on a FPGA Allowing Signed and Fraction Coefficients with Coefficients Obtained Using Remez Exchange Algorithm", *International Journal of Advancements in Technology*, Vol 1, No 2, October 2010, pp. 203-210.
- [8] Vijiya Prakash A.M. and K.S. Gurumurthy, "A Novel VLSI Architecture for Low Power FIR Filter", *International Journal of Advanced Engineering & Application*, January 2011, pp. 218-224.



# LIDAR - Light Detection And Ranging Technology

A. Prof. Vikram H. Doshi, B. Prof. Madhav K. Dave

**Abstract**— LIDAR is an acronym for Light Detection and Ranging, sometimes also referred to as Laser Altimetry or Airborne Laser Terrain Mapping (ALTM). The LiDAR system basically consists of integration of three technologies, namely, Inertial Navigation System (INS), LASER, and GPS. The Global Positioning System (GPS) has been fully operational for over a decade, and during this period, the technology has proved its potential in various application areas.

Geo Spatial Information is an important input for all planning and developmental activities especially in the present era of digital mapping and decision support systems. LiDAR is much faster than conventional photogrammetric technology and offers distinct advantage over photogrammetry in some application areas. Its development goes back to 1970s and 1980s, with the introduction of the early NASA-LiDAR systems, and other attempts in USA and Canada (Ackermann, 1999).

**Index Terms**—Agriculture, Applications, Archaeology, Astronomy, Atmospheric, Biology, Conservation, Design, Geology Science, LiDAR - Light Detection And Ranging, Meteorology, Micro pulse, Operation, Physics, Soil.

## I. INTRODUCTION

**L**IDAR (Light Detection And Ranging) is an optical remote sensing technology that measures properties of scattered light to find range and/or other information of a distant target. The prevalent method to determine distance to an object or surface is to use laser pulses. Like the similar radar technology, which uses radio waves, the range to an object is determined by measuring the time delay between transmission of a pulse and detection of the reflected signal. LIDAR technology has application in Geometrics, archaeology, geography, geology, geomorphology, seismology, forestry, remote sensing and atmospheric physics. Applications of LIDAR include **ALSM** (*Airborne Laser Swath Mapping*), **laser altimetry** or LIDAR Contour Mapping. The acronym **LADAR** (*Laser Detection and Ranging*) is often used in military contexts. The term "**laser radar**" is also in use even though LIDAR does not employ microwaves or radio waves, which is definitional to radar.

The primary difference between LIDAR and radar is that LIDAR uses much shorter wavelengths of the electromagnetic spectrum, typically in the ultraviolet, visible, or near infrared range. In general it is possible to image a feature or object only

about the same size as the wavelength, or larger. Thus lidar is highly sensitive to aerosols and cloud particles and has many applications in atmospheric research and meteorology.

An object needs to produce a dielectric discontinuity to reflect the transmitted wave. At radar (microwave or radio) frequencies, a metallic object produces a significant reflection. However non-metallic objects, such as rain and rocks produce weaker reflections and some materials may produce no detectable reflection at all, meaning some objects or features are effectively invisible at radar frequencies. This is especially true for very small objects (such as single molecules and aerosols).

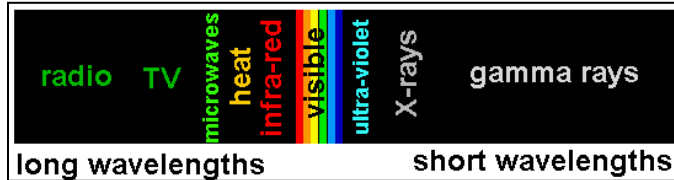
Lasers provide one solution to these problems. The beam densities and coherency are excellent. Moreover the wavelengths are much smaller than can be achieved with radio systems, and range from about 10 micrometers to the UV (ca. 250 nm). At such wavelengths, the waves are "reflected" very well from small objects. This type of reflection is called backscattering. Different types of scattering are used for different lidar applications, most common are Rayleigh scattering, Mie scattering and Raman scattering as well as fluorescence. Based on different kinds of backscattering, the LIDAR can be accordingly called Rayleigh LiDAR, Mie LiDAR, Raman LiDAR and Na/Fe/K Fluorescence LIDAR and so on. The wavelengths are ideal for making measurements of smoke and other airborne particles (aerosols), clouds, and air molecules.

A laser typically has a very narrow beam which allows the mapping of physical features with very high resolution compared with radar. In addition, many chemical compounds interact more strongly at visible wavelengths than at microwaves, resulting in a stronger image of these materials. Suitable combinations of lasers can allow for remote mapping of atmospheric contents by looking for wavelength-dependent changes in the intensity of the returned signal.

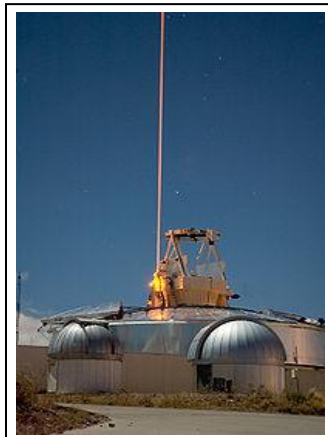
LIDAR has been used extensively for atmospheric research and meteorology. With the deployment of the GPS in the 1980s precision positioning of aircraft became possible. GPS based surveying technology has made airborne surveying and mapping applications possible and practical. Many have been developed, using downward-looking LIDAR instruments mounted in aircraft or satellites. A recent example is the NASA Experimental Advanced Research Lidar.

II. WHAT'S A LIDAR?

A lidar is similar to the more familiar *radar*, and can be thought of as a *laser radar*. In a radar, radio waves are transmitted into the atmosphere, which scatters some of the power back to the radar's receiver. A lidar also transmits and receives electromagnetic radiation, but at a higher frequency. Lidars operate in the **ultraviolet**, **visible** and **infrared** region of the electromagnetic spectrum.



Different types of physical processes in the atmosphere are related to different types of light scattering. Choosing different types of scattering processes allows atmospheric composition, temperature and wind to be measured.



**LIDAR** is an acronym which stands for **L**ight **D**etection **A**nd **R**anging (radar is also an acronym). A simplified block diagram of a lidar contains a transmitter, receiver and detector system. The lidar's transmitter is a

laser, while its receiver is an optical telescope.

A FASOR used at the Starfire Optical Range for LIDAR and laser guide star experiments is tuned to the sodium D2a line and used to excite sodium atoms in the upper atmosphere.

This lidar (laser range finder) may be used to scan buildings, rock formations, etc., to produce a 3D model. The lidar can aim its laser beam in a wide range: its head rotates horizontally, a mirror flips vertically. The laser beam is used to measure the distance to the first object on its path.



III. OPERATION

Different kinds of lasers are used depending on the power and wavelength required. The lasers may be both cw (continuous wave, on continuous like a light bulb) or pulsed (like a strobe light).

Gain mediums for the lasers include, gases (e.g. HeNe = Helium Neon or Xenon Fluoride), solid state diodes, dyes and

crystals (e.g. Nd:YAG = Neodymium:Yttrium Aluminum Garnet).

For some lidar applications more than 1 kind of laser is used. Here is the transmitter for Western's Purple Crow Lidar, which uses both cw and pulsed lasers. The final output of both channels of the transmitter is pulsed with a pulse repetition rate of 20 times per second and a pulse width of about 7 ns (1 ns = 1 x 10<sup>-9</sup>s).



The receiving system records the scattered light received by the receiver at fixed time intervals.

Lidar typically use extremely sensitive detectors called photomultiplier tubes to detect the backscattered light.

Photomultiplier tubes (shown below) convert the individual quanta of light, photons, first into electric currents and then into digital photocounts which can be stored and processed on a computer.

Amazing when you consider the electric currents generated are on the order of picoamps (1 pA = 10<sup>-12</sup> A; a 60 W light bulb draws a current of 0.5 A!).

The photocounts received are recorded for fixed time intervals during the return pulse.

The times are then converted to heights called range bins since the speed of light is well known.

So to find range bins from time:

$$rangebin = \Delta Z = (C * \Delta t) / 2$$

where *c* is the speed of light, 3 x 10<sup>8</sup> m/s.

So if each range bin is 160 ns long the height of each bin is 24 m. The range-gated photocounts are then stored and analyzed by a computer. Sources for information in this section can be found here.

IV. DESIGN

In general there are two kinds of lidar detection schema: "incoherent" or direct energy detection (which is principally an amplitude measurement) and Coherent detection (which is best for doppler, or phase sensitive measurements). Coherent systems generally use Optical heterodyne detection which being more sensitive than direct detection allows them to operate a much lower power but at the expense of more complex transceiver requirements.

In both coherent and incoherent LIDAR, there are two types of pulse models: **micropulse lidar** systems and **high energy** systems. Micropulse systems have developed as a result of the ever increasing amount of computer power available combined with advances in laser technology. They use considerably less energy in the laser, typically on the order of one microjoule, and are often "eye-safe," meaning they can be used without safety precautions. High-power systems are common in atmospheric research, where they are widely used for measuring many atmospheric parameters: the height, layering and densities of clouds, cloud particle properties (extinction coefficient, backscatter coefficient, depolarization), temperature, pressure, wind, humidity, trace gas concentration (ozone, methane, nitrous oxide, etc.).

There are several major components to a LIDAR system:

1. **Laser** — 600-1000 nm lasers are most common for non-scientific applications. They are inexpensive but since they can be focused and easily absorbed by the eye the maximum power is limited by the need to make them eye-safe. Eye-safety is often a requirement for most applications. A common alternative 1550 nm lasers are eye-safe at much higher power levels since this wavelength is not focused by the eye, but the detector technology is less advanced and so these wavelengths are generally used at longer ranges and lower accuracies. They are also used for military applications as 1550 nm is not visible in night vision goggles unlike the shorter 1000 nm infrared laser. Airborne topographic mapping lidars generally use 1064 nm diode pumped YAG lasers, while bathymetric systems generally use 532 nm frequency doubled diode pumped YAG lasers because 532 nm penetrates water with much less attenuation than does 1064 nm. Laser settings include the laser repetition rate (which controls the data collection speed). Pulse length is generally an attribute of the laser cavity length, the number of passes required through the gain material (YAG, YLF, etc.), and Q-switch speed. Better target resolution is achieved with shorter pulses, provided the LIDAR receiver detectors and electronics have sufficient bandwidth.
2. **Scanner and optics** — How fast images can be developed is also affected by the speed at which it can be scanned into the system. There are several options to scan the azimuth and elevation, including dual oscillating plane mirrors, a combination with a polygon mirror, a dual axis scanner. Optic choices affect the angular resolution and range that can be detected. A hole mirror or a beam splitter are options to collect a return signal.
3. **Photodetector and receiver electronics** — Two main photodetector technologies are used in lidars: solid state photodetectors, such as silicon avalanche photodiodes, or photomultipliers. The sensitivity of

the receiver is another parameter that has to be balanced in a LIDAR design.

4. **Position and navigation systems** — LIDAR sensors that are mounted on mobile platforms such as airplanes or satellites require instrumentation to determine the absolute position and orientation of the sensor. Such devices generally include a Global Positioning System receiver and an Inertial Measurement Unit (IMU).

## V. APPLICATIONS

This LIDAR - equipped mobile robot uses its LIDAR to construct a map and avoid obstacles. Other than those applications listed above, there are a wide variety of applications of LIDAR, as often mentioned in National LIDAR Dataset programs.



### A. Archaeology

LIDAR has many applications in the field of

archaeology including aiding in the planning of field campaigns, mapping features beneath forest canopy, and providing an overview of broad, continuous features that may be indistinguishable on the ground. LIDAR can also provide archaeologists with the ability to create high-resolution digital elevation models (DEMs) of archaeological sites that can reveal micro-topography that are otherwise hidden by vegetation. LiDAR-derived products can be easily integrated into a Geographic Information System (GIS) for analysis and interpretation. For example at Fort Beausejour - Fort Cumberland National Historic Site, Canada, previously undiscovered archaeological features have been mapped that are related to the siege of the Fort in 1755. Features that could not be distinguished on the ground or through aerial photography were identified by overlaying hillshades of the DEM created with artificial illumination from various angles. With LiDAR the ability to produce high-resolution datasets quickly and relatively cheaply can be an advantage. Beyond efficiency, its ability to penetrate forest canopy has led to the discovery of features that were not distinguishable through traditional geo-spatial methods and are difficult to reach through field surveys.

### B. Meteorology and Atmospheric Environment

The first LIDAR systems were used for studies of atmospheric composition, structure, clouds, and aerosols. Initially based on ruby lasers, LIDAR for meteorological applications was constructed shortly after the invention of the laser and represent one of the first applications of laser technology.

*Elastic backscatter LIDAR* is the simplest type of lidar and is typically used for studies of aerosols and clouds. The backscattered wavelength is identical to the transmitted wavelength, and the magnitude of the received signal at a given range depends on the backscatter coefficient of scatterers at that range and the extinction coefficients of the scatterers along the path to that range. The extinction coefficient is typically the quantity of interest.

*Differential Absorption LIDAR (DIAL)* is used for range-resolved measurements of a particular gas in the atmosphere, such as ozone, carbon dioxide, or water vapor. The LIDAR transmits two wavelengths: an "on-line" wavelength that is absorbed by the gas of interest and an off-line wavelength that is not absorbed. The differential absorption between the two wavelengths is a measure of the concentration of the gas as a function of range. DIAL LIDARs are essentially dual-wavelength elastic backscatter LIDARs.

*Raman LIDAR* is also used for measuring the concentration of atmospheric gases, but can also be used to retrieve aerosol parameters as well. Raman LIDAR exploits inelastic scattering to single out the gas of interest from all other atmospheric constituents. A small portion of the energy of the transmitted light is deposited in the gas during the scattering process, which shifts the scattered light to a longer wavelength by an amount that is unique to the species of interest. The higher the concentration of the gas, the stronger the magnitude of the backscattered signal.

*Doppler LIDAR* is used to measure wind speed along the beam by measuring the frequency shift of the backscattered light. Scanning LIDARs, such as NASA's HARLIE LIDAR, have been used to measure atmospheric wind velocity in a large three dimensional cone. ESA's wind mission ADM-Aeolus will be equipped with a Doppler LIDAR system in order to provide global measurements of vertical wind profiles. A doppler LIDAR system was used in the 2008 Summer Olympics to measure wind fields during the yacht competition. Doppler LIDAR systems are also now beginning to be successfully applied in the renewable energy sector to acquire wind speed, turbulence, wind veer and wind shear data. Both pulsed and continuous wave systems are being used. Pulsed systems using signal timing to obtain vertical distance resolution, whereas continuous wave systems rely on detector focusing.

*Synthetic Array LIDAR* allows imaging LIDAR without the need for an array detector. It can be used for imaging Doppler velocimetry, ultra-fast frame rate (MHz) imaging, as well as for speckle reduction in coherent LIDAR.

### C. Wind power

Lidar is sometimes used on wind farms to more accurately measure wind speeds and wind turbulence, and an experimental lidar is mounted on a wind turbine rotor to measure oncoming horizontal winds, and proactively adjust blades to protect components and increase power.

### D. Geology and Soil Science

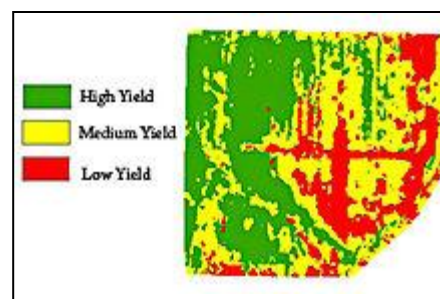
High-resolution digital elevation maps generated by airborne and stationary LIDAR have led to significant advances in geomorphology, the branch of geoscience concerned with the origin and evolution of Earth's surface topography. LIDAR's abilities to detect subtle topographic features such as river terraces and river channel banks, measure the land surface elevation beneath the vegetation canopy, better resolve spatial derivatives of elevation, and detect elevation changes between repeat surveys have enabled many novel studies of the physical and chemical processes that shape landscapes. In addition to LIDAR data collected by private companies, academic consortia have been created to support the collection, processing and archiving of research-grade, publicly available LIDAR datasets. The National Center for Airborne Laser Mapping (NCALM), supported by the National Science Foundation, collects and distributes LIDAR data in support of scientific research and education in a variety of fields, particularly geoscience and ecology.

In geophysics and tectonics, a combination of aircraft-based LIDAR and GPS have evolved into an important tool for detecting faults and measuring uplift. The output of the two technologies can produce extremely accurate elevation models for terrain that can even measure ground elevation through trees. This combination was used most famously to find the location of the Seattle Fault in Washington, USA. This combination is also being used to measure uplift at Mt. St. Helens by using data from before and after the 2004 uplift. Airborne LIDAR systems monitor glaciers and have the ability to detect subtle amounts of growth or decline. A satellite based system is NASA's ICESat which includes a LIDAR system for this purpose. NASA's Airborne Topographic Mapper is also used extensively to monitor glaciers and perform coastal change analysis. The combination is also used by soil scientists while creating a soil survey. The detailed terrain modeling allows soil scientists to see slope changes and landform breaks which indicate patterns in soil spatial relationships.

### E. Agriculture

Agricultural Research Service scientists have developed a way to incorporate LIDAR with yield rates on agricultural fields. This technology will help farmers direct their resources toward the high-yield sections of their land.

LIDAR also can be used to help farmers determine which areas of their fields to apply costly fertilizer. LIDAR can create a topological map of the fields and reveals the slopes



and sun exposure of the farm land. Researchers at the Agricultural Research Service blended this topological information with the farm land's yield results from previous years.

From this information, researchers categorized the farm land into high-, medium-, or low-yield zones. This technology is valuable to farmers because it indicates which areas to apply the expensive fertilizers to achieve the highest crop yield.

#### F. Physics and astronomy

A worldwide network of observatories uses lidars to measure the distance to reflectors placed on the moon, allowing the moon's position to be measured with mm precision and tests of general relativity to be done. MOLA, the Mars Orbiting Laser Altimeter, used a LIDAR instrument in a Mars-orbiting satellite (the NASA Mars Global Surveyor) to produce a spectacularly precise global topographic survey of the red planet.

In September, 2008, NASA's Phoenix Lander used LIDAR to detect snow in the atmosphere of Mars.

In atmospheric physics, LIDAR is used as a remote detection instrument to measure densities of certain constituents of the middle and upper atmosphere, such as potassium, sodium, or molecular nitrogen and oxygen. These measurements can be used to calculate temperatures. LIDAR can also be used to measure wind speed and to provide information about vertical distribution of the aerosol particles.

At the JET nuclear fusion research facility, in the UK near Abingdon, Oxfordshire, LIDAR Thomson Scattering is used to determine Electron Density and Temperature profiles of the plasma.

#### G. Biology and conservation

LIDAR has also found many applications in forestry. Canopy heights, biomass measurements, and leaf area can all be studied using airborne LIDAR systems. Similarly, LIDAR is also used by many industries, including Energy and Railroad, and the Department of Transportation as a faster way of surveying. Topographic maps can also be generated readily from LIDAR, including for recreational use such as in the production of orienteering maps.

In oceanography, LiDAR is used for estimation of phytoplankton fluorescence and generally biomass in the surface layers of the ocean. Another application is airborne lidar bathymetry of sea areas too shallow for hydrographic vessels.

In addition, the Save-the-Redwoods League is undertaking a project to map the tall redwoods on California's northern coast. LIDAR allows research scientists to not only measure the height of previously unmapped trees but to determine the biodiversity of the redwood forest. Stephen Sillett who is working with the League on the North Coast LIDAR project claims this technology will be useful in directing future efforts to preserve and protect ancient redwood trees.

#### H. Military and law enforcement

One situation where LIDAR has notable non-scientific application is in traffic speed enforcement, for vehicle speed measurement, as a technology alternative to radar guns. The technology for this application is small enough to be mounted in a hand held camera "gun" and permits a particular vehicle's speed to be determined from a stream of traffic. Unlike RADAR which relies on doppler shifts to directly measure speed, police lidar relies



Police officer using a hand-held LIDAR speed gun

on the principle of time-of-flight to calculate speed. The equivalent radar based systems are often not able to isolate particular vehicles from the traffic stream. LIDAR has the distinct advantage of being able to pick out one vehicle in a cluttered traffic situation as long as the operator is aware of the limitations imposed by the range and beam divergence.

LIDAR does not suffer from "sweep" error when the operator uses the equipment correctly and when the LIDAR unit is equipped with algorithms that are able to detect when this has occurred. A combination of signal strength monitoring, receive gate timing, target position prediction and pre-filtering of the received signal wavelength prevents this from occurring. Should the beam illuminate sections of the vehicle with different reflectivity or the aspect of the vehicle changes during measurement that causes the received signal strength to be changed then the LIDAR unit will reject the measurement thereby producing speed readings of high integrity. For LIDAR units to be used in law enforcement applications a rigorous approval procedure is usually completed before deployment. The use of many reflections and an averaging technique in the speed measurement process increase the integrity of the speed reading. Vehicles are usually equipped with a horizontally oriented registration plate that, when illuminated, causes a high integrity reflection to be returned to the LIDAR - despite the shape of the vehicle. In locations that do not require that a front or rear registration plate is fitted, headlamps and rear-reflectors provide almost ideal retro-reflective surfaces overcoming the reflections from uneven or non-compliant reflective surfaces thereby eliminating "sweep" error. It is these mechanisms which cause concern that LIDAR is somehow unreliable.

Most traffic LIDAR systems send out a stream of approximately 100 pulses over the span of three-tenths of a second. A "black box" proprietary statistical algorithm picks and chooses which progressively shorter reflections to retain from the pulses over the short fraction of a second.

Military applications are not yet known to be in place and are possibly classified, but a considerable amount of research is underway in their use for imaging. Higher resolution systems collect enough detail to identify targets, such as tanks. Here the name LADAR is more common.

Utilizing LIDAR and THz interferometry wide area raman spectroscopy, it is possible to detect chemical, nuclear, or biological threats at a great distance. Further investigations regarding long distance and wide area spectroscopy are currently conducted by Sandia National Laboratories.

Five LIDAR units produced by the German company Sick AG were used for short range detection on Stanley, the autonomous car that won the 2005 DARPA Grand Challenge.

A robotic Boeing AH-6 performed a fully autonomous flight in June 2010, including avoiding obstacles using LIDAR.

### I. Vehicles

LIDAR has been used in Adaptive Cruise Control (ACC) systems for automobiles. Systems such as those by Siemens and Hella use a lidar device mounted on the front of the vehicle, such as the bumper, to monitor the distance between the vehicle and any vehicle in front of it. In the event the vehicle in front slows down or is too close, the ACC applies the brakes to slow the vehicle. When the road ahead is clear, the ACC allows the vehicle to accelerate to a speed preset by the driver.

### J. Imaging

3-D imaging is done with both scanning and non-scanning systems. "3-D gated viewing laser radar" is a non-scanning laser radar system that applies the so-called gated viewing technique. The gated viewing technique applies a pulsed laser and a fast gated camera. There are ongoing military research programmes in Sweden, Denmark, the USA and the UK with 3-D gated viewing imaging at several kilometers range with a range resolution and accuracy better than ten centimeters.

Coherent Imaging LIDAR is possible using Synthetic array heterodyne detection which is a form of Optical heterodyne detection that enables a staring single element receiver to act as though it were an imaging array. This avoids the need for a gated camera and all ranges from all pixels are simultaneously available in the image.

Imaging LIDAR can also be performed using arrays of high speed detectors and modulation sensitive detectors arrays typically built on single chips using CMOS and hybrid CMOS/CCD fabrication techniques. In these devices each pixel performs some local processing such as demodulation or gating at high speed down converting the signals to video rate so that the array may be read like a camera. Using this technique many thousands of pixels / channels may be acquired simultaneously. In practical systems the limitation is light budget rather than parallel acquisition.

LIDAR has been used in the recording of a music video without cameras. The video for the song "House of Cards" by Radiohead is believed to be the first use of real-time 3D laser scanning to record a music video.

### K. Aerial Surveying - 3D mapping

Aerial LiDAR surveying from a paraplane operated by Scandinavian Laser Surveying.

Airborne LIDAR sensors are used by companies in the Remote Sensing field to create point clouds of the earth ground for further processing (e.g. used in forestry). A common format for saving these points (with parameters like x, y, return, intensity, elevation) is the LAS file format.



## VI. ADVANTAGES

There are several advantages of LIDAR data. First, it is a very versatile technology that has been used for atmospheric studies, bathymetric surveys, glacial ice investigations, and numerous other applications. It is finding a lot of use in terrain mapping. Here we see that this technology is a very cost effective method of terrain data collection. It offers high precision and high point density data for DTM modeling. Moreover, it has been shown to accelerate the project schedule, upwards to 30% because the DTM data processing can begin almost immediately. It is, theoretically, not restricted to daylight nor cloud cover like aerial photography, although if aerial imagery is being collected simultaneously, as it is commonly done, then those limitation will affect the particular project. In coastal zones and forest areas, LIDAR is considered as a superior data collection tool over conventional photogrammetric techniques where it is extremely difficult to locate terrain points in the imagery. LIDAR requires only one opening through a tree canopy to "see" the ground whereas photogrammetry requires that the same ground point be visible from two exposure stations. This would cut down on the amount of area identified as "obscured terrain" on a contour map.

Photogrammetry is a mature science that is still undergoing technological advances. The products derived from this mapping system are well received and the limitations are understood. While LIDAR appears to be an excellent alternative to photogrammetric mapping, there are several disadvantages to LIDAR when the two technologies are compared.

## VII. DISADVANTAGES

There are several disadvantages as well. While the data collection appears to be cost competitive, the upfront cost of equipment acquisition is very significant, on the order of \$1 million. This could be a hard sell since amortization would have to be spread over a very short period since the technology, like that of computers, will probably experience a lot of change over the next two to three years. That is a lot of imagery to collect over a short period of time. While LIDAR is an active system that can be, theoretically, used 24 hours a day, it cannot be used above cloud cover or when fog, smoke, mist, rain, or snow storms are present. Additionally, high winds and turbulence will cause problems with the inertial system.

There are problems with data collected over water, which leads to suspect delineation of water boundaries using LIDAR by itself. LIDAR systems are not capable of determining break lines. Laser scan data are collected in a more or less regular spacing pattern. In other words, it cannot be pointed on a specific feature. For example, a 2 meter-wide ditch may not be shown on a LIDAR dataset with a spacing of 5 meters. Thus, LIDAR data are often augmented with break line data compiled from photogrammetric methods.

Being a relatively new technology, standards have not been established that could help guide the user as to the quality of the results. There are a number of efforts underway to alleviate this problem.

When elevation data are compiled from photogrammetric processes, the operator has a "cartographic license" when selecting points for measurement. Contour lines are generally smoothed to reflect the actual representation of the terrain. A large boulder, as an example, may be a ground surface point captured by LIDAR and the resulting data may depict this as a high point in the terrain. The photogrammetric operator would not use this point in the data collection process as a terrain point for DEM generation.

## VIII. CONCLUSION

It is clearly evident that many within the GIS industry are looking at LIDAR as an economical and accurate means of collecting both feature and terrain data. Indeed, this technology is growing. Like any new technological tool, there are times when the technology is misused. Just like GPS has not made conventional terrestrial surveying obsolete, LIDAR will not soon supplant photogrammetric mapping as an economical and accurate method of collecting data about features on the earth. As the technology matures, as new data processing techniques are developed, and as standards are

developed, it is safe to say that LIDAR will become an important data collection methodology available to the user community.

## REFERENCES

- [1] Experimental Advanced Research Lidar', NASA.org. Retrieved 8 August 2007.
- [2] [http://www.sciencegl.com/gis\\_dem/index.html](http://www.sciencegl.com/gis_dem/index.html)
- [3] An Airborne Altimetric LiDAR tutorial
- [4] Mikkelsen, Torben & Hansen, Kasper Hjorth et al. Lidar wind speed measurements from a rotating spinner Danish Research Database & Danish Technical University, 20 April 2010. Retrieved: 25 April 2010.
- [5] <http://en.wikipedia.org/wiki/LIDAR>
- [6] <http://airborneimagingusa.com/>
- [7] Ackermann, F., Airborne laser scanning present status and future expectations. ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 54, 1999.
- [8] Councillor Quarterly, Summer 2007 Volume 6 Issue 3
- [9] <http://superlidar.colorado.edu/Classes/Lidar2011/LidarLecture14.pdf>
- [10] Amzajerdian, Farzin; Pierrottet, Diego F.; Petway, Larry B.; Hines, Glenn D.; Roback, Vincent E.. "Lidar Systems for Precision Navigation and Safe Landing on Planetary Bodies"
- [11] Medina, Antonio. *Three Dimensional Camera and Rangefinder*. **January 1992**. United States Patent 5081530.
- [12] NASA. 'NASA Mars Lander Sees Falling Snow, Soil Data Suggest Liquid Past' *NASA.gov* (29 September 2008). Retrieved 9 November 2008.
- [13] Tom Paulson. 'LIDAR shows where earthquake risks are highest', Seattle Post, April 18, 2001.
- [14] [http://www.arete.com/index.php?view=stil\\_mcm](http://www.arete.com/index.php?view=stil_mcm)
- [15] <http://www.ino.ca/en-CA/Achievements/Description/project-p/short-range-bioaerosol-threat-detection.html>

# Design and Implementation of GUI and Operation of Excimer Laser using LabVIEW™

A. Sandeep Joshi

**Abstract--** The paper describes the development of GUI based control software for control/operation, maintenance and data logging of Coherent CompexPro® 102 Excimer Laser (ArF, 193 nm) using LabVIEW™ instrument control software.

**Index-Terms —** LabVIEW™, Excimer Laser, Graphical user interface, Coherent, RS-232, control.

## I. INTRODUCTION

The Coherent COMPexPro102® excimer laser (Fig.1) is an ArF gas based UV laser operating at a wavelength of 193 nm. This laser uses Ar, F2, Ne, and He gases for its operation. The laser can deliver a maximum energy of ~ 200 mJ at 15 ns pulse width and capable of operating at a pulse repetition rate of 1-20 Hz. The laser is normally controlled using a tethered control panel which has a cable length restriction of around 3 meters. In order that the laser can be controlled and monitored remotely, it is essential that a GUI based interface be developed. The block-diagram of system showing the gas feed system, the I/O system and the energy meters is shown in Fig.2.



Fig.1. Physical layout of the excimer laser system

Currently, the operation of the laser is carried out manually using the wired keypad. The main motivation of the development of this GUI is for easier user convenience as well as for future use when the laser will be used with the plasma accelerator system, where the laser operation will be carried remotely and the laser system will be integrated with the accelerator system operations *via* LabView™ instrument control software.

Routine laser operation (setting of parameters and running of

A. pursuing M.Tech (Instrumentation) at the Nirma University, Ahmedabad (sandeepbm2100@gmail.com).

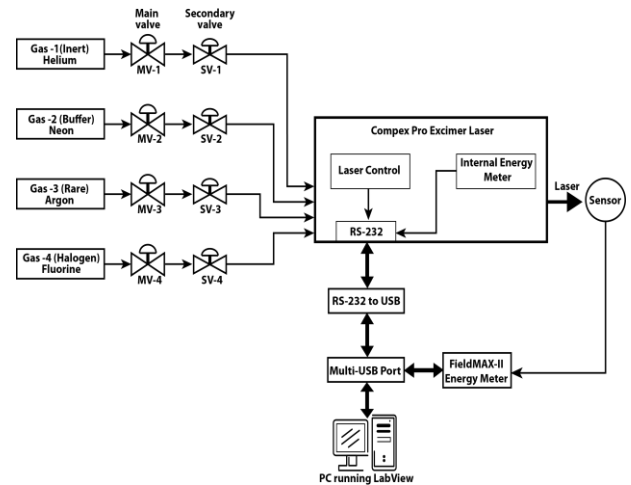
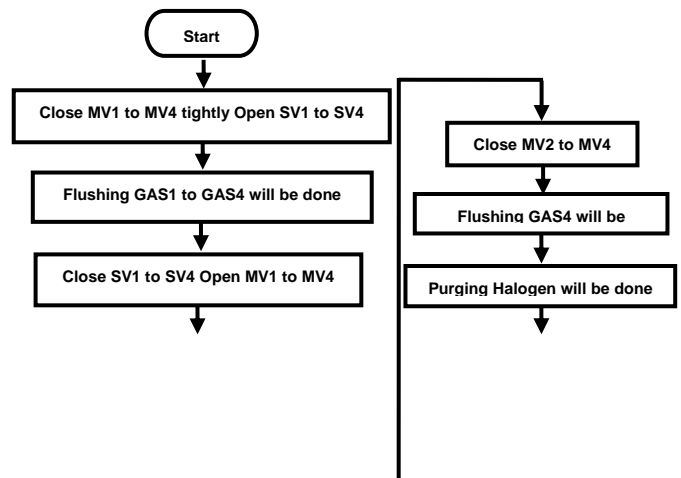


Fig.2.Schematic diagram of the excimer laser system showing the gas lines, I/O interfaces and energy calibration layout.

laser), maintenance operations like gas fill and energy calibration, logging of laser parameters etc., which are currently being carried out manually will be automated using the GUI.

This requires development of the essential LabVIEW™ virtual instrument modules, GUI interfaces and the required display screens for laser control, parameter display and documentation of the operational parameters of the laser system. The aim of this paper is develop a completely automated module for calibration of the internal energy meter of the excimer laser with an external laser energy meter (Coherent FieldMax-II). This one-switch, automated operation provides a linear-fitted calibration of the internal energy meter with that of a pre-calibrated commercial energy meter.

The I/O control of the laser is established via a standard RS-232 interface coupled to the PC through a RS-232-to-USB interface. LabVIEW™ tools like web-publishing Tool, structure palette, string, VISA palette have been used to design this graphical user interface. The laser is designed to be controlled either from the control panel or from a remote PC but not from both at the same time.





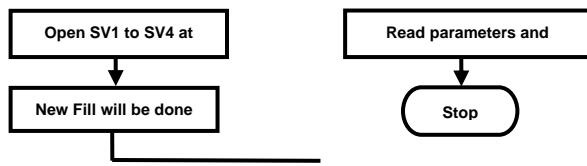


Fig.3 Flowchart of Gas Fill Operation

Field-MAX-II-P energy meter with an energy sensor optimized and calibrated for use with 193 nm UV laser is used for external energy measurement. The UV sensor incorporates a DUV quartz diffuser for increased coating damage resistance. The energy meter has USB I/O connectivity and hence could be seamlessly integrated into the laser GUI.

II. MODULE FOR GAS-FILL OPERATION

Excimer Laser uses a mixture of various gases depending on the wavelength of operation of the laser. The 193 nm operation of the laser uses he, Ne, Ar and F gases in various proportions which are decided by the laser system at the time of gas filling operation. The gas fills have to be carried out periodically so as to maintain the laser at its operating energy. The flow-chart of gas-filling process based on which the GUI has been designed is shown in Fig.3. In this GUI, gas flushing, new-fill, purge-gas line and data logging are automatically carried out.

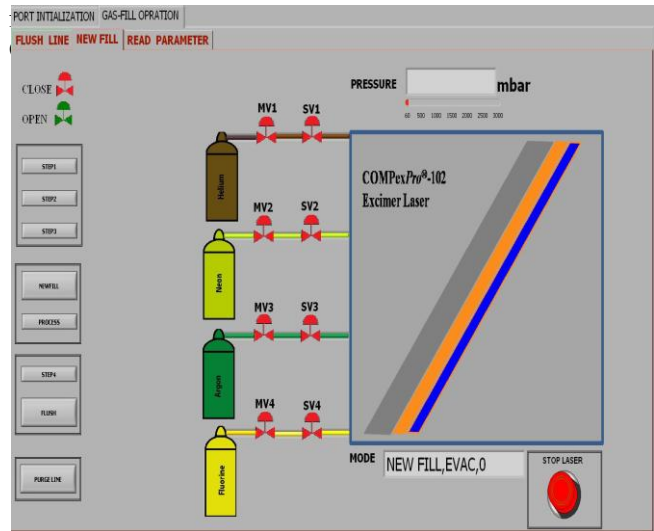


Fig.5 Front-Panel for New fill and Purge line. In future, the whole operation could be fully automated if the manual valves are replaced by electrically controlled valves. Once the gas fill is over, the GUI will automatically flush the gas lines and purge the lines with inert gas for safety. The GUI for the gas fill is shown in Fig.5.

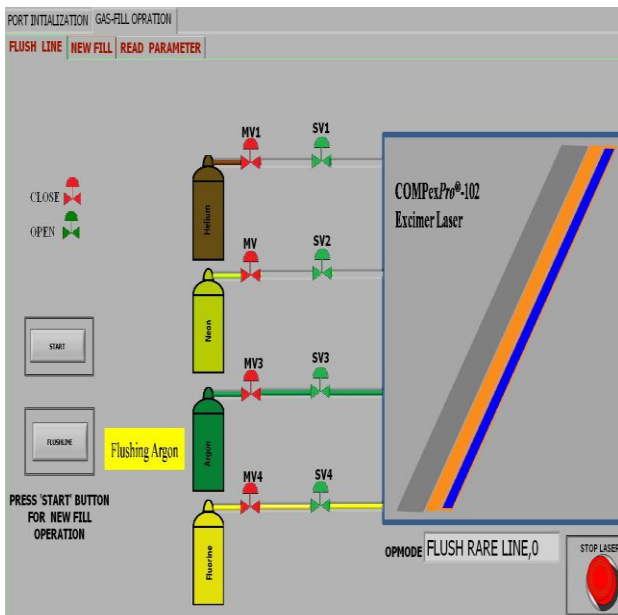


Fig.4 Front-Panel for Flushing gas lines.

A. Front-panel for flushing gas lines

Gas flushing is an essential step to clear the gas lines of all impurities before the actual gas fill operation is initiated. When this operation is initiated from the “Flush Line” GUI, all the gas lines will be flushed using helium gas and since Fluorine is a corrosive gas, that line is

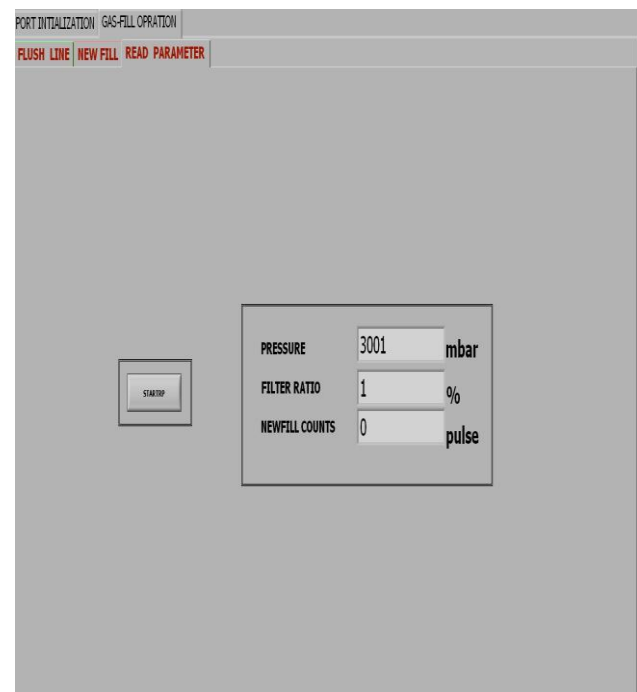
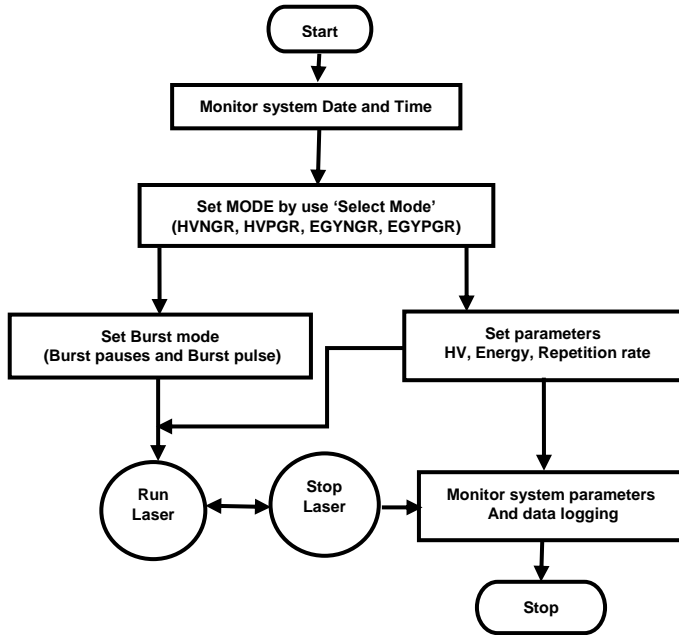


Fig.6 Front-Panel for post gas fill parameters.



C. Front panel for display of laser maintenance parameters

Other system maintenance parameters such as gas pressures, filter ratio (lifetime of the halogen filter), total laser pulses, pulses after refill, tube temperature etc can be displayed. By using the “STARTRP” command, these parameters can be logged into an Excel file for laser maintenance purposes. The flow chart for this GUI is shown in Fig.6.

### III. MODULE FOR MONITOR AND SET PARAMETER

#### FOR EXCIMER LASER

This GUI is basically to display the various operating and set parameters of the laser. Parameters such as laser energy, laser repetition rate and mode of operation (Constant energy or constant voltage) can be set. Also, there is an option to activate the burst-mode operation of the laser if desired so by the user. Flowchart for this GUI is shown in Fig.7. Front-panel for this seen in Fig.8.

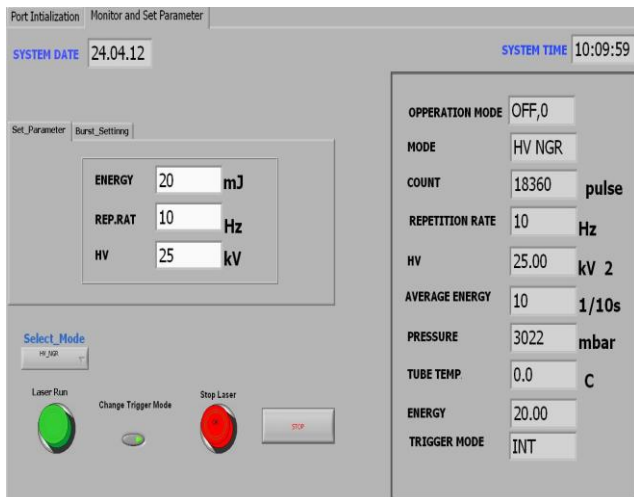
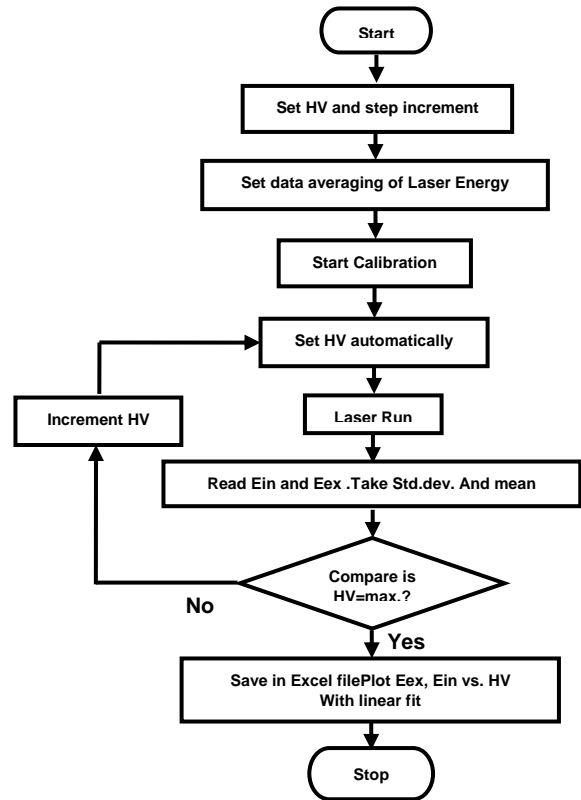


Fig.8. Front-panel of Monitor and Set parameter



There is four basic operating modes under “Select Mode” viz, HV-PGR, HV-NGR (constant voltage mode with and without gas purge), EGY-PGR, EGY-NGR, (constant energy mode with and without gas purge) for setting the running mode for the excimer laser. At a time only one of the four options can be activated and can be changed only after laser is off.

### IV. MODULE FOR ENERGY CALIBRATION

The internal energy meter of the excimer laser taps a small percentage of the laser beam to monitor the beam energy. This is used for stabilizing the laser output. Over time, the calibration of this internal energy meter deteriorates and needs to be corrected over time. The calibration of the internal energy meter can be done against a pre-calibrated external energy meter. By plotting the outputs of both the internal and external energy meters at various operating voltages, one can determine if the internal energy meter is correct. If the calibration is found to off by more

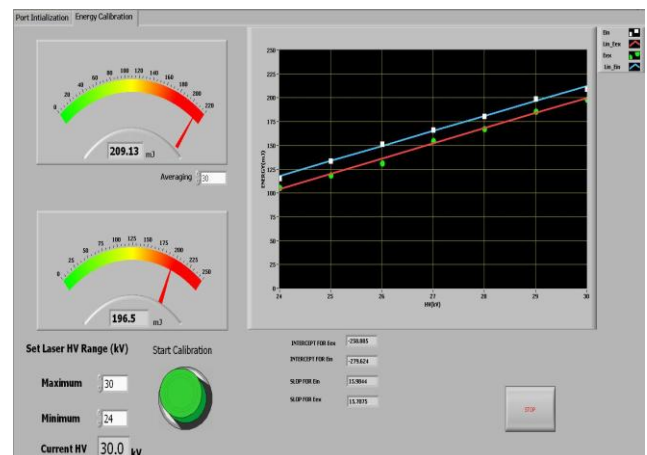


Fig.10 Front-Panel of Energy-Calibration

than an acceptable percentage, then the internal energy meter needs be physically recalibrated.

The flow-chart for energy calibration is shown in Fig.9. The “START CALIBERATION” process requires the outputs of both the laser and the external energy meter to be read, data averaged and standard deviation determined for various laser voltage

Time	HV	Ein	Linfit Eex	Eex	Linfit Ein	Slope Ein	Slope Eex	STD Eex
(hr:min:sec)	(kV)	(mJ)	(mJ)	(mJ)	(mJ)			
11:13:52	24	103.13	104.18	109.43	108.85	14.65	14.53	7.35
11:15:50	25	123.2	118.83	113.64	123.39	14.65	14.53	8.15
11:17:32	26	137.42	133.48	129.25	137.92	14.65	14.53	7.20
11:19:46	27	155.79	148.12	151.34	152.45	14.65	14.53	7.45
11:21:52	28	168.85	162.77	163.17	166.99	14.65	14.53	6.43
11:23:43	29	181.06	177.42	178.28	181.52	14.65	14.53	7.97
11:25:52	30	193.73	192.07	191.75	196.06	14.65	14.53	7.35

Fig.11. Data for energy calibration saved in excel file.

levels. The averaged data is plotted in the graph *in-situ*. After the process ends, the data is saved in Excel form, and finally plotted against laser high voltage, as shown in Fig 10. The data points are then linearly fitted to obtain slopes, which are then compared to determine if the calibration of the internal energy meter is within acceptable limits. The internal statistical math modules of LabVIEW™ are used for this. The logged data in Excel format is shown in Fig.11.

V. LABVIEW™ DRIVERS FOR EXCIMER LASER

In this Virtual Instrument driver (VI) file, all the Sub-VI drivers for different commands related to the excimer laser operation have been developed and incorporated. The VI-Tree is schematically shown in Fig.12.

A. Flush Command

There are four “Flush” commands are used to flush gas lines of the various gases used for operation of the excimer laser (Halogen, Buffer, Rare, an Inert gases).

B. Write Command

The four sub-routines under this VI are set/control functions related to running of the laser *ie.*, “Run Laser”, “Stop Laser”, “Change value of HV”, “Repetition Rate” .

C. Read Command

All parameters (like operation-mode, mode, repetition-rate, HV, Energy, Pressure, System time and date, tube temperature, etc.) of excimer laser are read by running this VI.

D. Logical Sub-routine

There are five sub-routines under this VI. Password sub VI is used for the security purpose. Call VI is used for to call VI from

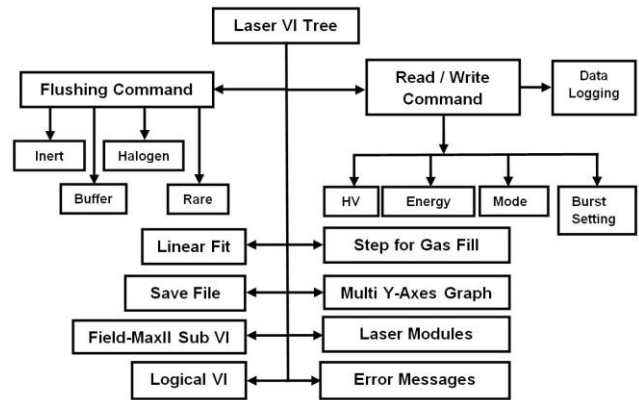


Fig.12.Excimer Laser VI Tree

folder where it has been saved. Regression is used to find relation between two parameters by linear fit.

E. Mathematics

These sub-VI are used for to find mean, min, max, standard deviation and slope for data in the calibration operation.

F. Save file

Data logging is done by this VI. Here, all parameters of laser are saved in excel file with time and date.

G. Multi Y-axes Graph

Multi X-Y graphs are plotted when this VI is run. It plots two parameters at Y-axes and one parameter on X-axes.

H. Step for Gas-fill

All messages related to step wise explanation of the manual part of the gas-fill operation are displayed to the user when this VI is run.

I. Modules for Excimer Laser

There are four sub VI modules (like Port Initialization, Gas-Fill operation, Monitor and set parameter and Energy calibration) for excimer laser which have been designed under this module.

VI. CONCLUSIONS

GUI for port initialization, setting and monitor all parameters of excimer laser, gas fill operation and energy calibration for excimer laser are designed by serial interface using RS-232, DB-9 connector. But, all parameters of excimer laser are monitor and set with one second or two second delay. Virtual instrument of Excimer laser and external energy meter has been designed using LabVIEW™ software. Real-Time Data acquisition will be done in various file like .doc,.pdf,.xls,.txt etc. All laser operation will be done on internet

in LabVIEW™ environment if all computers are connected in LAN. Automatic programming will be design for Laser operations and it will be change. Laser operation will be become faster and safe than it operate manually. Laser operations are easy to operate and understand by user. Laser Data will be transferred for long distance using RS-232 to RS-485 converter.

#### ACKNOWLEDGMENT

I would like to thank Dr. Ravi A.V. Kumar and Mr. Ankit Sharma for all their timely and useful guidance and help which resulted in the successful completion of this project.

I would also like to thank Mr. Yogesh Yeole for helping me overcome many problems that I faced while using LabVIEW™. I would also like to thank Mr. Mohandas, Ms. Kanchan Mahavar, Mr. Abhijeet and Mr. Jinto Thomas of IPR for help rendered on various occasions right through this project work.

#### VII. REFERENCES

- [1] John Park, Steve Mackay, and Edwin Wright, "Practical Data Communications for Instrumentation and Control," 2003.
- [2] NATIONAL INSTRUMENT LabVIEW™ 2010 –software.
- [3] JOVITHA JEROME, "VIRTUAL INSTRUMENTATION USING Lab VIEW™," Published by Asoke K. Ghosh., 2009.
- [4] Hexin RS-485 to RS-S232 interface converter Model-485.
- [5] COHERENT INTERFACING MANUAL COMPEXpro (RoHS).
- [6] User Manual FieldMaxII-P Laser Energy Meter.
- [7] Bella G Liptak, "Process Software and Digital networking" 3<sup>rd</sup> edition.
- [8] User Manual Energy max Sensors.

# Effect of Scaling & Parameter Variation on Design of CMOS Voltage Comparator

A. Gireeja Amin, B. Rajanikant Soni, C. N. M. Devashrayee, D. Amisha Naik

**Abstract**— Higher speed and higher density are the main thrusts of CMOS technology and are achieved by device miniaturization. In deep submicron geometries, the supply voltage is scaled down to prevent reliability hazards such as oxide breakdown and hot carrier effects and also to reduce energy per operation of digital circuits. Lowering the supply voltage directly reduces the signal swing, which in turn makes the design of high-speed wide dynamic range mixed-signal circuits a challenge. This paper demonstrates effect of scaling and effect of parameter variations on the design of a High Speed CMOS Comparator. This paper describes the pre layout simulation results of Differential Comparator which is simulated using Eldo spice simulator for TSMC 0.35 $\mu$ m and 0.13 $\mu$ m technology.

**Index Terms**— Comparator, Low Voltage, High Speed, Power supply, Scaling.

## I. INTRODUCTION

The CMOS is becoming the major processing technology for implementing digital signal processing systems. With the rapid advancement in digital circuit technologies, very high speed and large data handling capability is already available. However to realize of single chip analog or digital system, analog circuits with high operating speed and greater accuracy are required. The A/D converter is one of the key interface blocks between the continuous time domain and the discrete time domain. Its performance is highly dependent on the voltage comparator. Comparator is widely used in the process of converting analog signals to digital signals. In high-speed analog-to-digital converters, comparator design has a crucial influence on the overall performance that can be achieved.

Higher speed and higher density are the main benefits of CMOS VLSI technology and are achieved by shrinking the feature size of the MOSFET devices. In the past, Constant Voltage (CV) scaling has been used for CMOS (down to mm) to achieve higher performance and to maintain supply voltage compatibility. However, as MOSFET miniaturization reaches to deep submicron sizes (0.5 $\mu$ m and below), supply voltage scaling based on Quasi-Constant Voltage (QCV) scaling must be adopted to assure reliability [Kakumu90]. Hot

carrier effects and gate oxide breakdown are the two important reliability factors determining how high power supply voltage can be. The Semiconductor Industry Association (SIA) roadmap predicts a supply voltage of 0.9V

for semiconductors by the year 2010 [SIA94]. The move toward lower supply voltages is also fueled by low-power battery powered portable devices [Thomas93]. An ideal power supply for a battery powered system is a single-cell ( 1.2V-0.9V ) off-the-shelf battery.

Year	1995	1998	2001	2004	2007	2010
Feature size ( $\mu$ m)	0.35	0.25	0.18	0.13	0.1	0.07
PS(V)	3.3	2.5	1.8	1.5	1.2	0.9

Table 1: SIA semiconductor technology roadmap

Integrated circuits are moving increasingly into the mixed-signal world where more parts of a system are implemented on a single chip. Performance and density of digital circuits are expected to follow the famous Moore's law—achieving a 2X increase in circuit density and a slower rate of (1.4-1.2)X increase in circuit speed every three years. Energy per operation in digital circuits ( $E=CV^2$ ) is also reduced due to supply voltage scaling in deep submicron CMOS. However, lowering the supply voltage directly reduces the signal swing in analog circuits, which in turn makes the design of wide dynamic range mixed signal circuits a challenge. In deep submicron CMOS, the charge carriers velocity will be saturated at the pinch-off point in the channel and as a result the classical quadratic saturation current becomes  $I_{dsat} = (V_{GS} - V_{th})^\alpha$ , where  $\alpha = 1.2$  to 1.3 instead of 2 [Hu94]. Therefore, 1.4X circuit speed improvement every generation will slow down to about 1.2X. The effect of technology scaling on various design constrains is shown in table:2 .

CMOS Technology Scaling		Design Constraints			
		Channel length	Oxide	$V_{DSAT}$	Circuit Complexity
Intrinsic Speed	$f_T \uparrow$	Short	Thin	Large	Low
Power Supply	$V_{DD} \downarrow$	/	/	Small	Low
Thermal Noise	$4kT(\gamma \cdot g_m) \uparrow$	Long	Thick	/	Low
Intrinsic Gain	$g_m r_{out} \downarrow$	Long	Thick	Small	High
Device Matching	$W, L \downarrow$	Long	/	/	Low
Device Modeling	SCE	Long	Thick	/	/

Table 2: Impact of technology scaling on analog circuit design

Parameter variations encompass a range of variation types, including process variations due to manufacturing phenomena, voltage variations due to manufacturing and runtime phenomena, and temperature variations due to varying activity levels and power dissipations—in fact, these three main sources are often referred to as PVT (process-voltage-temperature) variations. Process variations are static

and manifest themselves as die-to-die (D2D), within die (WID) variations, and wafer-to-wafer variations (W2W), while temperature and voltage variations are dynamic phenomena. Temperature variations stem from different activity factors among cores, functional units, from different circuit structures, and from non uniformities exacerbated by activity-dependent IR drops. These are exacerbated by temperature-dependent leakage-current variations (i.e., varying the I term) or switching activity that causes voltage droops due to circuit inductance and possibly insufficient decoupling capacitances. These three variation sources exhibit a number of feedback loops.

Process variations affect leakage, which affects both voltage and temperature. Temperature then affects leakage forming a feedback loop between the two parameters [1]. Parameter variations cause chip characteristics to deviate from the uniform, ideal values desired at design time. Three major sources of variation are often discussed: process variations, which consist of deviations in the manufactured properties of the chip, such as feature size, dopant density, etc.; voltage variations due to non uniform power-supply distribution, switching activity, and IR drop : and temperature variations due to non-uniformities in heat flux of different functional units under different workloads as well as the impact of non-uniformities in the chip’s interface to its package.

The concern is that as technology scales PVT variations will increase in severity, resulting in worse required design margins.

Section II describes the functioning of the CMOS comparator, section III describes the Prelayout Simulation of the CMOS Comparator in different technologies and its effect on it. Section IV describes the effect of parameter variation. Section V Results and conclusion.

II. COMPARATOR

Comparators often provide a link between the analog and digital domains. The sampled signal of analog input is applied to combination of comparators to determine the digital equivalent of analog signal. The comparator is a circuit that compares that compares analog signal with another analog signal or reference signal and outputs a binary signal based on comparison.

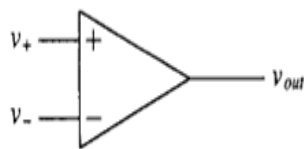


Figure 1 Symbol of Comparator

The high speed comparator consists of three stages; the input preamplifier, a positive feedback or decision stage, and an output buffer.

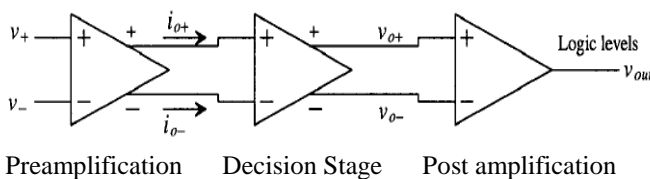


Figure 2 Block Diagram of Comparator

Differential amplifier will act as a preamplifier circuit. Higher gain can be obtained if the size of M3 and M4 is larger than the size of M31 and M41. Latch stage is also called decision circuit and capable of discriminating mV difference in the logic level. A self-biasing differential amplifier is used as the comparator output buffer. The main purpose of the output buffer is to convert the output of the decision circuit into a logic signal. Current Mirror is used in the design for providing a constant current over a wide range of voltage.

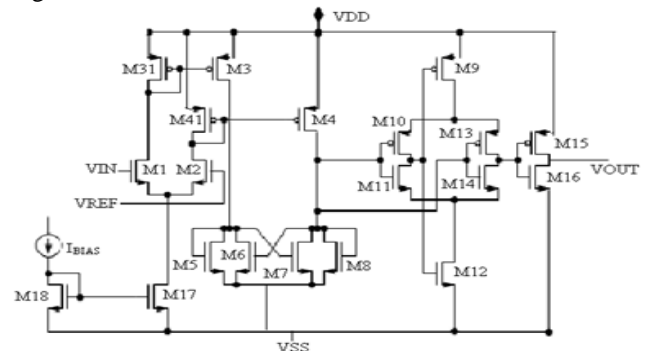


Figure 3 Schematic of CMOS Comparator

TECHNOLOGY	0.35µm	0.13µm
Supply Voltages VDD VSS	±1.8 V	±1.5 V
ICMR	-1.8V to 1.8V	-1.5V to 1.5V
Propagation Delay	<2 ns	<2 ns
Max Operating Frequency	100MHz	100MHz
Power Dissipation	15 mW	15mW

Table-3:Design Specification

III. PRELAYOUT SIMULATIONS

Prelayout simulation is carried out using the ELDO Spice in TSMC 0.13µm. The parameter used is BSIM3v3 Level – 53. In this section static characteristics of comparator like offset, ICMR, Frequency Response and dynamic characteristics propagation delay is observed from the simulation.

Propagation Delay is the amount of time between the time when VP - VN = 0 and the output is 50% between initial and final value. From the simulation results shown in the figure10 and 11 we observed the propagation delay is 420.39ps and 765.65ps for 0.13 µm and 0.35 µm technology respectively

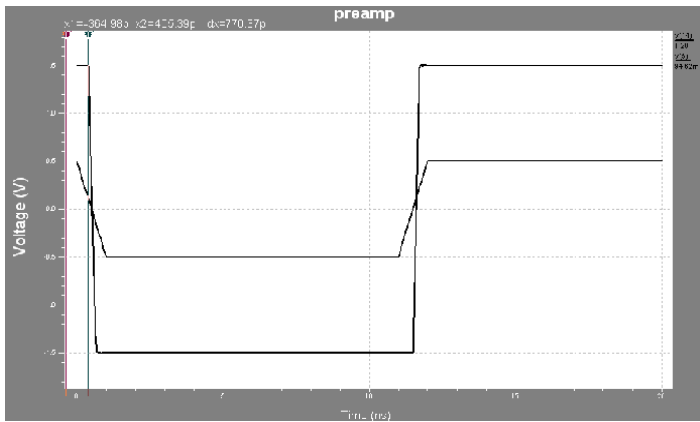


Figure 4 Transient Analysis in 0.13 μm

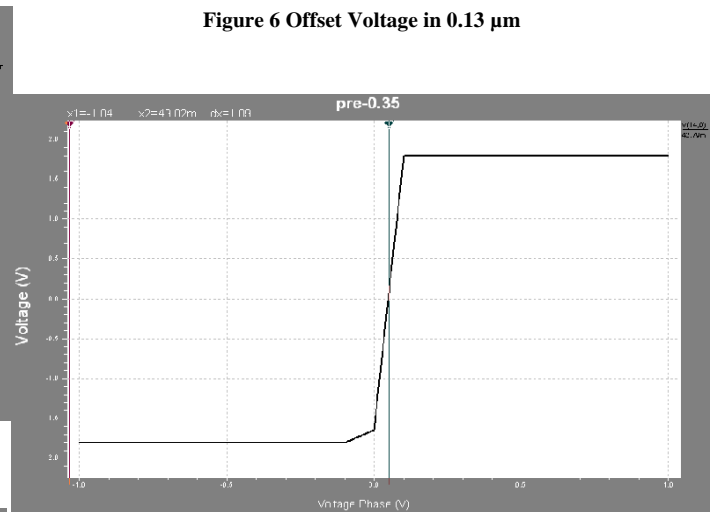


Figure 6 Offset Voltage in 0.13 μm

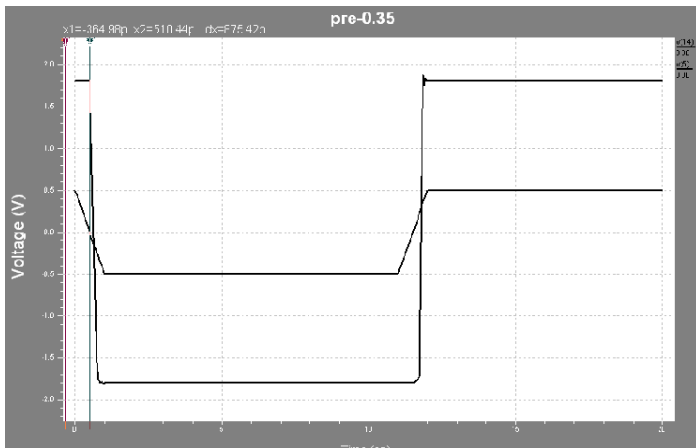


Figure 5 Transient Analysis in 0.35 μm

Figure 7 Offset Voltage in 0.35 μm

ICMR is the input voltage range over which the comparator function normally. The input voltage range over which the comparator can detect  $V_P = V_N$ . From the simulation results shown in the figure 8 and 9 we observed the ICMR is -1V to 1.5V and -1.4V to 1.8V for 0.13 μm and 0.35 μm technology respectively.

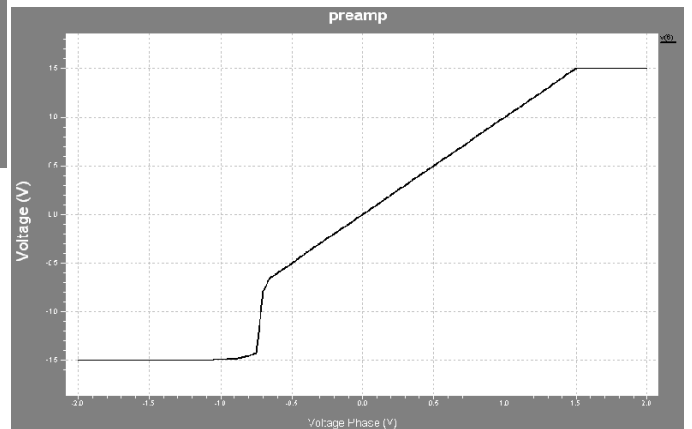


Figure 8 ICMR in 0.13 μm

Offset voltage is nothing but distance between the origin and the output.[3] Here the input voltage from -0.5V to 0.5V for respective technology. From the simulation results shown in the figure 6 and 7 we observed the offset voltage is 52.53mv and 49.18mv for 0.13 μm and 0.35 μm technology respectively

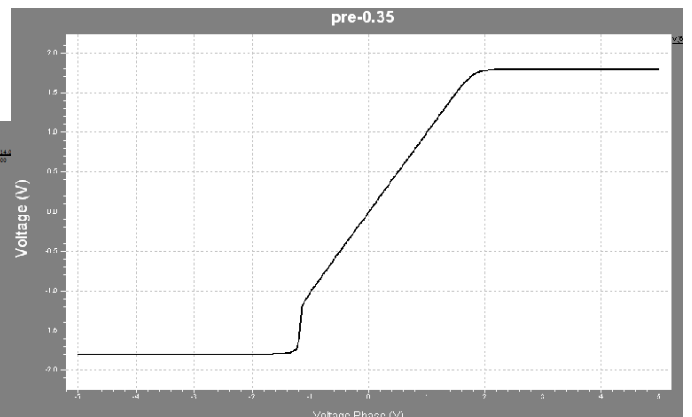
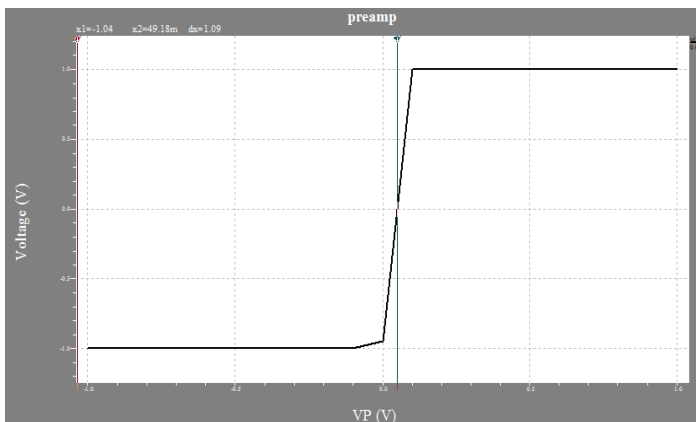


Figure 9 ICMR in 0.35 μm



From the simulation results shown in figure 10 and 11 we observed the gain of the comparator is 74dB and 78 dB and

operating frequency of 306Mhz and 82.75Mhz for 0.13  $\mu\text{m}$  and 0.35  $\mu\text{m}$  technology respectively.

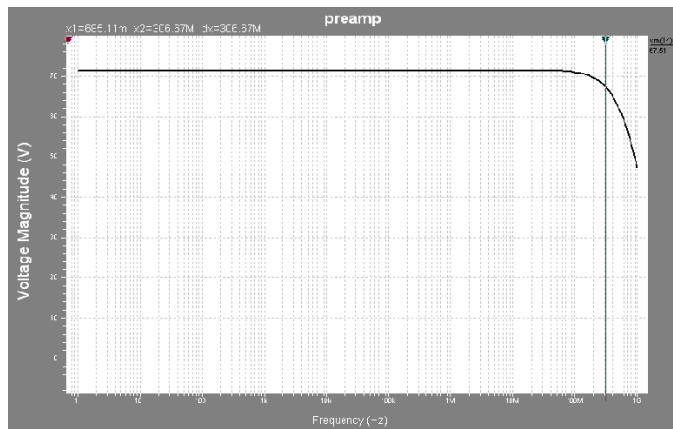


Figure 10 Frequency Response in 0.13  $\mu\text{m}$

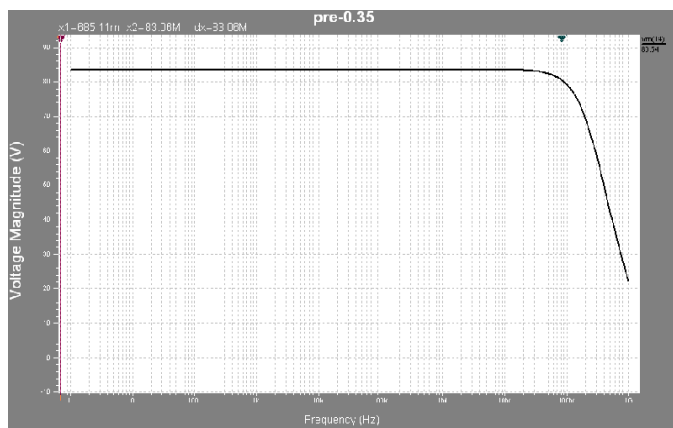


Figure 11 Frequency Response in 0.35  $\mu\text{m}$

IV. EFFECT OF PARAMETER VARIATION ON CMOS COMPARATOR

Supply voltage will have more influence on the power dissipation of the design hence on power consumption. By reducing the supply voltage we can reduce the power dissipation of the design. So we have observed the effect of variation in the supply voltage on power and gain of the design.

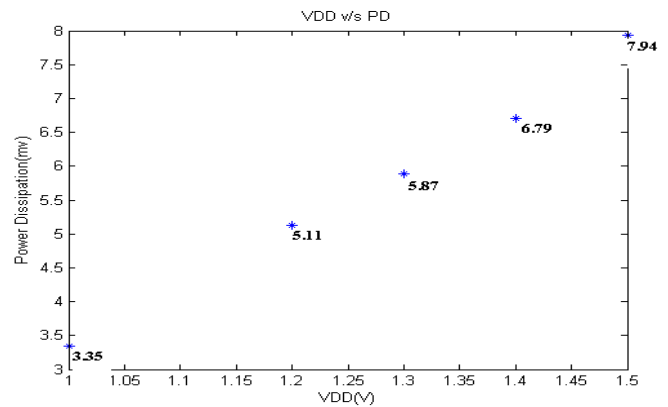


Figure 14 Supply Voltage v/s Power Dissipation

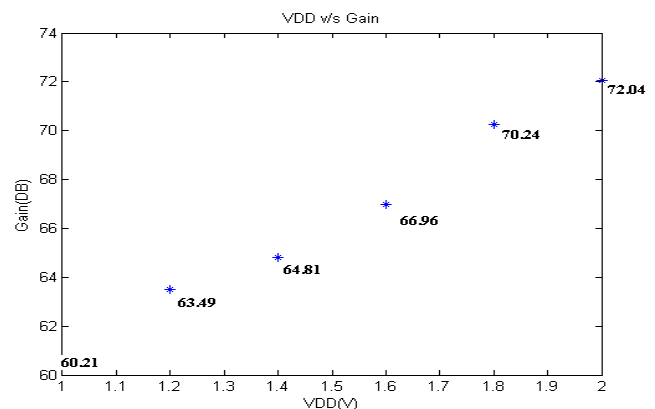


Figure 15 Supply Voltage v/s Gain

TECHNOLOGY	0.35 $\mu\text{m}$	0.13 $\mu\text{m}$
Supply Voltages VDD VSS	$\pm 1.8\text{ V}$	$\pm 1.5\text{ V}$
Offset	49.18mV	52.53mV
ICMR	-2.2 V to 2.2 V	-1.5 V to 2. 2 V
Propagation Delay	765.67ps	420.39ps
Gain	78dB	74dB
Power	11.65mW	7.84mW
Speed	82.75Mhz	306Mhz

Table-5 effect of scaling on cmos comparator

TECHNOLOGY	0.13 $\mu\text{m}$	0.13 $\mu\text{m}$
Supply Voltages VDD VSS	$\pm 1.0\text{ V}$	$\pm 1.5\text{ V}$
Offset	52.53mV	52.53mV
ICMR	-0.5 V to 1.0 V	-1.5 V to 2. 2 V
Propagation Delay	475ps	420.39ps
Gain	74dB	74dB
Power Dissipation	3.35mW	7.84mW
Speed	107.78Mhz	306Mhz

Table-5 effect of power supply variation on cmos comparator



## V. CONCLUSION

The Pre amplifier based comparator is designed using 0.13um and 0.35um Technology to see the effect of parameter variation in terms of VDD and effect of scalling on speed,power,offset,delay and gain.As seen from the results increase in power supply gives improvement in gain bandwidth product at the cost of increase in power dissipation as well as as we scall down the technology speed inceses at the cose of increase in offset voltage.

## REFERENCES

- [1] Eric Humenay, David Tarjan, Kevin Skadron” Impact of Parameter Variations on MultiCore Chips”33<sup>rd</sup>International symposium on computer architecture,ASGI’06,2006.
- [2] P. Allen and D. Holberg, CMOS Analog Circuit Design, Oxford University Press, 2002.
- [3] R. Jacob Baker Harry W. Li David E. Boyce. CMOS Circuit Design, Layout and Simulation. IEEE Press Series on Microelectronics Systems, 2005.
- [4] B. Razavi, Design of Analog CMOS Integrated Circuits, McGraw-Hill, 2001.
- [5] Behzad Razavi, Principles of Data Conversion System Design, IEEE Press, 1995, 189~190.
- [6] Song Ye, Jie Wu, “*An Ultra-high-speed Comparator for ADC in 90nm CMOS Technology*”, IEEE Press,2009 ,978-1-4244-4076-4
- [7] *Mustafa Parlak ve Yajar Giirbiiz,* Low Power, Variable Supply CMOS Comparator “, IEEE Press, 2004,0-7803-83 I8-4104
- [8] Jiang Li, Xu Weisheng, Yu Youling “A high-speed and high-resolution CMOS comparator with three-stage preamplifier” Journal of Semiconductors, Vol. 31, No. 4 April 2010

# Effective Radio Channel Modeling: Concept and Technique

A. Nirali A. Kotak

**Abstract:** This paper presents the critical comparative analysis of efficient wireless channels using the modeling of wireless communication system. In the last few years, the telecommunication industries' development has focused on an intensive use of broadband systems, which are characterized by high quality features. For this issue, new technologies with lower Bit Error Rate abilities have been designed. That is the very base of the wireless communication system that allows a fast deployment as well as efficient channel modeling. As wireless communication is in great demand for different personal communication, effective channel modeling becomes imperative. The paper helps to understand the concept of effective modeling of a wireless system with the proper use of various types of wireless channels under appropriate circumstances.

In this paper, we first build up a wireless communication simulator including digital modulation, OFDM modulator and Demodulator, different channel models such as AWGN, Rayleigh and Rician fading channels. In the first phase the modeling and simulation of wireless communication system has been presented along with different types of channels. In the second phase the comparative analysis of AWGN, Rayleigh and Rician channel has been done under practical scenario by taking the various values of Bit Error Rate and S/N ratio as well as Doppler shift. The whole model includes the basic mechanism of OFDM system so as to have effective communication by saving the required value of bandwidth.

**Key Words:** AWGN Channel, Rayleigh Channel, Rician Channel, BER, SNR and OFDM.

## I. INTRODUCTION

Mobile communications and wireless network have experienced massive growth and commercial success in the recent years. However, the radio channels in mobile radio systems are usually not amiable as the wired one.

Unlike wired channels that are stationary and predictable, wireless channels are extremely random and time-variant. It is well known that the wireless multi-path channel causes an arbitrary time dispersion, attenuation, and phase shift, known as fading, in the received signal. Fading is caused by interference between two or more versions of the transmitted signal which arrived at the receiver at slightly different times.

In general, the needs for communication channel are providing full mobility, capable of higher data rate, better integrated interoperable with wide range of equipment, easier to configure and use. The performance of wideband wireless systems, like, cellular mobile radio, indoors-wireless communications, and personal communication services is limited by multipath fading.

The major complexity in the multipath environment results from tracking each multipath at the receiver. Therefore, effective modeling and diverting the multiple paths to the receiver is extremely important for the performance of the overall systems. In this paper, three broadly classified channel models are explained. For wireless communication system different parameters such as modulation technique, bandwidth, frequency range, bit rate etc in some sense or the other depends up on the channel modeling. Therefore, it is very important to look forward for effective channel modeling. The paper is divided into two parts. First part deals with the modeling and simulation of different channels and second part is regarding their comparative analysis.

## II SYSTEM MODELING USING AWGN CHANNEL

The simplest type of channel is the Gaussian channel. It is often referred to the additive white Gaussian noise (AWGN) channel. Basically, it is the noise generated in the receiver side if we assume that the transmitter is ideal and noiseless. This type of noise is assumed to have a constant power spectral density over the whole channel bandwidth and its amplitude probability density function (PDF) obeys the statistics of a Gaussian distribution.

Gaussian noise is very important in the analysis of communication system performance. The classical AWGN channel is always considered as the starting point to develop basic systems performance. Also, according to central limit theorem, even when there are a larger number of non-Gaussian independent noise sources, the mobile channel noise may still be approximated with a Gaussian distribution. This feature allows for simpler analysis of the communication system.

The traditional model for wireless communication system is built on QPSK i.e 4-QAM modulation scheme. The modeling setup includes Matlab R2007a, Simulink7 and Communications Block set 3 running on Windows XP SP2. Matlab Simulink includes all the mandatory function blocks as specified by the standard documents. The Model itself consists of three main components namely transmitter, receiver and channel. Transmitter and receiver components consist of modulation/demodulation blocks and OFDM modulation and demodulation blocks whereas channel is modeled as AWGN.

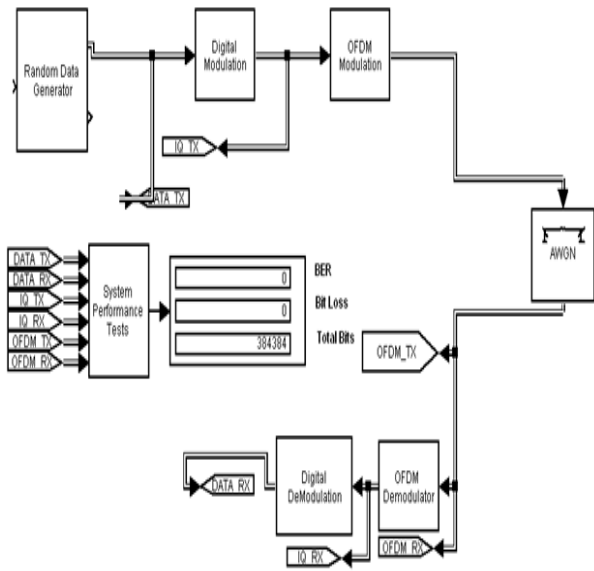


Figure 1 System model with AWGN channel

demodulator comprise of mainly IFFT and FFT blocks respectively for the orthogonal component generation.

Finally the processed data from the transmitter block is transmitted through AWGN channel which will give the different values of BER at various S/N level in the long distance communication. At higher db values of the SNR, the realization of AWGN channel gives the maximum accurate throughput about the signal variations because through the modeling of AWGN channel, the point to point variation in BER can be found out wrt SNR at longer distances.

Based on the model presented in this paper, and tests carried out, the performance was established based on 10 million symbols in each case. The performance is displayed in the above figures in terms of the magnitude spectrum, time-scatter plot and BER with reference to total transmitted bits for the output from the transmitter and FFT scope diagram for the transmitted signal.

### III SYSTEM MODELING USING RAYLEIGH CHANNEL

In a mobile radio communication system, one of the most devastating phenomena is fading. Fading is the direct result of multi-path propagation where radio waves propagate along different paths before arriving at the receiver antenna. These radio waves may arrive at receiver after different delays, with different amplitudes, and with different phases. Because there are so many different received signal components, constructive and destructive interference results in fading. This sort of channel is called a multi-path fading channel or Rayleigh channel. In urban areas where generally there is no line of sight component present, the analysis of the mobile radio channel can be obtained by modeling the Rayleigh channel rather than AWGN channel for mobile communication because the multipath structure can completely predicated through the modeling of this.

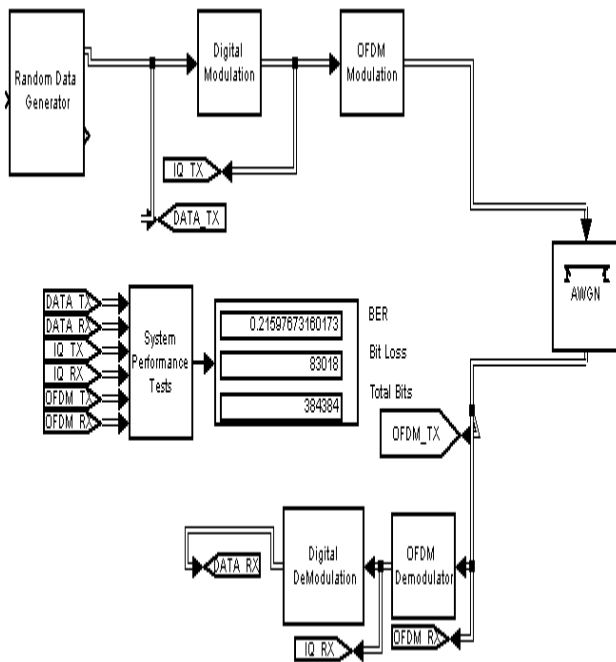


Figure 2 System model with AWGN channel

Randomizer is the block that generates the stream of data randomly which is processed further by the upcoming stages of the model. It operates on bit by bit basis. The randomly generated stream of bit pattern is modulated digitally by the QAM modulator.

Now before transmitting this modulated data stream through the channel, it is given to the OFDM modulator block which performs the multiplexing over the data in such a manner that approximately 50% of the bandwidth can be saved because of the mechanism of bifurcating the signal in in-phase and quadrature phase components. Basically OFDM modulator and

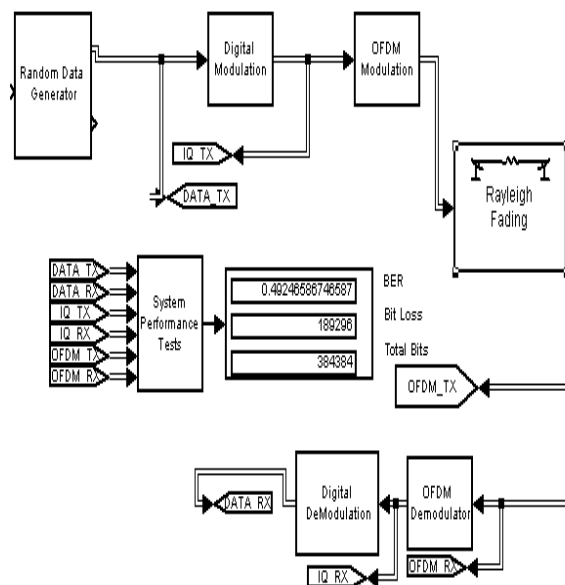


Figure 3 System model with Rayleigh channel

IV SYSTEM MODELING USING RICIAN CHANNEL

Rayleigh channel model is suitable for modeling urban areas that are characterized by many obstructions where a line of sight path does not exist between the transmitter and receiver. In suburban areas, a line of sight path may exist between the transmitter and receiver and this will give rise to Rician fading. Rician fading is characterized by a factor, which is expressed as the power ratio of the specular (los or dominant path) component to the diffused components. This ratio,  $k$ , defines how near to Rayleigh statistics the channel is. In fact when  $k = \infty$ , there is no fading at all and when  $k = 0$ , this means to have Rayleigh fading. The ratio is expressed linearly, not in decibels. While the Average path gain vector parameter controls the overall gain through the channel, the **K-factor** parameter controls the gain's partition into line-of-sight and diffuses components.

We can specify the **K-factor** parameter as a scalar or a vector. If the **K-factor** parameter is a scalar, then the first discrete path of the channel is a Rician fading process (it contains a line-of-sight component) with the specified **K-factor**, while the remaining discrete paths indicate independent Rayleigh fading processes (with no line-of-sight component). If the **K-factor** parameter is a vector of the same size as **Discrete path delay vector**, then each discrete path is a Rician fading process with a **K-factor** given by the corresponding element of the vector.

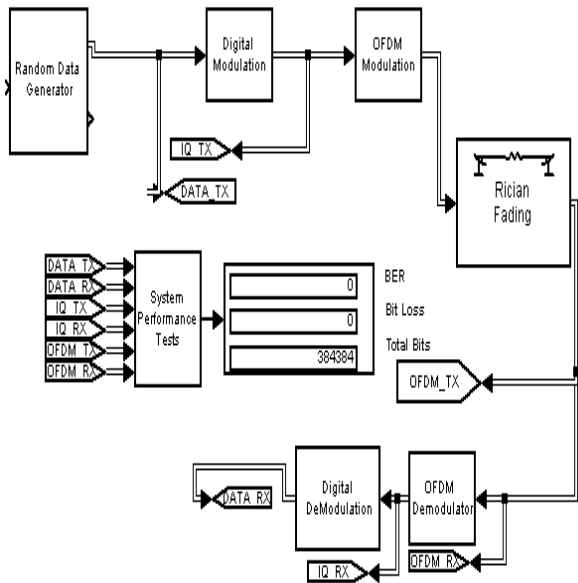


Figure 4 System model with Rician channel

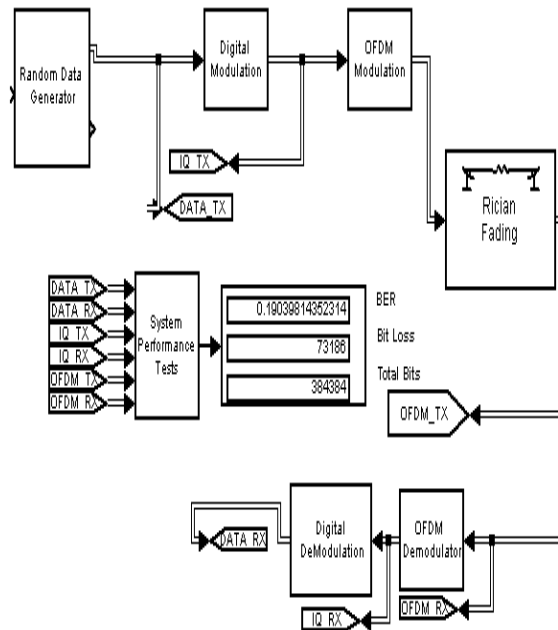


Figure 5 System model with Rician channel

V. PERFORMANCE EVALUATION AND CONCLUSION

Based on the simulation models presented in this paper and the tests carried out, the comparative performance analysis of the three wireless channels i.e. AWGN, Rayleigh and Rician can be described based on various parameters. For the long distance communication, at the higher value of SNR, it is appropriate to model wireless channel as AWGN channel because through this modeling we can analyze the variations in SNR through which the BER can be calculated. While for the small scale propagation, under the multi path environment, the modeling of wireless channel as a Rayleigh channel will give the accurate analysis if no line of sight component is present. But if surrounding atmosphere supports the single line of sight component along with the multipath structure then the Rician channel realization will give the appropriate modeling scenario for the modeling of any wireless system.

VII. FUTURE WORK

This simulation model of traditional wireless communication system is the fundamental of any modern emerging 4G or 5G wireless technology. For example, with the addition of multiple number of blocks such as MAC layer realization, source and channel coding in the present model, the simulation model of nowadays most prominent technology i.e. WiMAX can be obtained.

REFERENCES

[1] W. C. Y. Lee, "Overview of Cellular CDMA", *IEEE Trans. Veh. Techn.*, Vol. 40, No. 2, May 1991, pp. 291-302.  
 [2] D. L. Schilling et al., "Broad CDMA for Personal Communications Systems", *IEEE Comm. Mag.*, Vol. 29, No. 11, Nov. 1991, pp. 86-93.  
 [3] T. S. Rappaport, *Cellular Radio & Personal Communications: A Book of Selected Readings*, IEEE Press, New Jersey, 1995.

- 4] H. Hashemi, " The Indoor Radio Propagation Channel", *Proc. IEEE*, Vol. 81, No. 7, July 1993, pp. 941-968.
- 5] R. Ganesch and K. Pahlavan, "Statistical modeling and computer simulation of indoor radio channel", *IEE Proc. I*, Vol. 138, June 1991, pp. 153-161.
- [6] H. Suzuki, "A statistical model for urban mutipath propagation", *IEEE Trans. Commun*, Vol. 25, No. 7, July 1977, pp. 673-680.
- [7] T. S. Rappaport, S. Y. Seidel, and T. Takamizawa, "Statistical model impulse response model for factory and open plan building radio communication system design", *IEEE Trans. Comm.*, Vol. 39, No.5, May 1991, pp. 794-807. *Symposium "Wireless Trends in the 21<sup>st</sup> Century"*, Long Island, NY, Nov. 28-29,1995.
- [8] S. Y. Seidel and T. S. Rappaport, "Site-specific propagation prediction for wireless in-building personal communication system [17] S. J. Howard and K. Pahlavan, "Measurement and analysis of the indoor radio Channel", *IEEE Trans. Instr. Meas.*, Vol. 39, 1990, pp. 751-755.
- [9] 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 Mijsi and Vojislav B. Mistic, *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] Abdulrahman Yarali and Saifur Rahman, "WiMAX Broadband Wireless Access Technology: Services, Architecture and Deployment Models", *IEEE Transaction*

# Design Simulation and characterization of op amp based 3 Bit R-2R ladder DACs

A. Rajanikant Soni, B. Gireeja Amin, c. Dr. N.M.Devashrayee, d. Dr.Usha Mehta

**Abstract** - This paper presents a study on a digitally calibrated DAC, based on a strictly R-2R topology with operational amplifier which is able to derive high resolution - high performance DACs, in terms of INL and DNL. It has been proven by simulations that the performance of the conventional R-2R DAC can be optimized, regardless of resistors tolerance and the DAC resolution.

**Keywords**--- Digital to analog conversion, resolution, linearity, INL, DNL, glitch.

## I. INTRODUCTION

Digital to Analog Conversion performance is mainly characterized by its resolution, linearity and speed. Additional implementation characteristics include area and power dissipation. This paper presents a DAC architecture based on the conventional R-2R ladder topology that is able to derive a high-resolution, high-linearity and high-speed DAC. Probably the most popular digital-to-analog converter application is the digital audio compact disc player. Here digital information stored on the CD is converted into music via high precision DACs.

Here an  $N$ -bit digital word is mapped into a single analog voltage. Typically, the output of the DAC is a voltage that is some fraction of a reference voltage or current, such that

$$V_{OUT} = FV_{REF}$$

Where  $V_{out}$  is the analog voltage output,  $V_{ref}$  is the reference voltage, and  $F$  is the fraction defined by the input word,  $D$ , that is  $N$  bits wide. The number of input combinations represented by the input word  $D$  is related to the number of bits in the word by Number of input combinations =  $2^N$  A 3-bit DAC has a total of  $2^3$  or 8 total input values.

A converter with 3-bit resolution is able to map a change in the analog output which is equal to 1 part in 16. The maximum analog output voltage for any DAC must be limited by the value of some reference voltage  $V_{REF}$ .

If the input is an  $N$ -bit word, then the value of the fraction,  $F$ , can be determined by,

$$F = D/2^N$$

The *least significant bit (LSB)* refers to the rightmost bit in the digital input word. The LSB defines the smallest possible change in the analog output voltage. The LSB will always be denoted as  $D_0$ . One LSB can be defined as

$$1 \text{ LSB} = V_{REF}/2^n$$

The *most significant bit (MSB)* refers to the leftmost bit of the digital word,  $D$ . Generalizing to the  $N$ -bit DAC, the MSB would be denoted as  $DN-1$ .

Here the maximum analog output voltage that can be generated is known as *full-scale voltage*,  $V_{FS}$  and can be generalized to any  $N$ -bit DAC as

$$V_{FS} = \left( \frac{2^n - 1}{2^n} \right) V_{ref}$$

Full scale voltage is the highest analog output voltages.

### 1. Differential Nonlinearity

Non ideal components cause the analog increments to differ from their ideal values. The difference between the ideal and non ideal values is known as *differential Non linearity*, or  $DNL$  and is defined as,

$DNL_n = \text{Actual increment height of transition } n - \text{Ideal increment height}$

### 2. Integral Nonlinearity

Another important static characteristic of DACs is called *integral nonlinearity (INL)*.

$INL$  is defined as the difference between the data converter output values and a reference straight line drawn through the first and last output values, INL defines the linearity of the overall transfer curve and can be described as

$INL = \text{Output value for input code } n - \text{Output value of the reference line at that point}$

### 3. Mismatch Error related to DAC

The accuracy of the resistor string is obviously related to matching between the resistors, which ultimately determines the INL and DNL for the entire DAC. Suppose that the  $i$ -th resistor,  $R_j$ , has a mismatch error associated with it so that

$$R_i = R + \Delta R_i$$

Where  $R$  is the ideal value of the resistor and  $\Delta R_i$  is the mismatch error. Also suppose that the mismatches were symmetrical about the string so that the sum of all the mismatch terms were zero, or

$$= \sum_{i=1}^{2^n} \Delta R_i = 0 \quad \text{For } i = 0, 1, 2, \dots, 2^n - 1$$

However, including the mismatch, the actual value of the  $i$ -th voltage will be the sum of all the resistances up to and including resistor  $i$ , divided by the sum of all the resistances in the string, this can be represented by,

$$V_i = V_{ref} \frac{\sum_{k=1}^i R_k}{\sum_{k=1}^{2^n} R_k} = \frac{\sum_{k=1}^i R + \Delta R}{2^n R}$$

or finally, the value of the voltage at the  $i$ -th tap is

$$V_i = V_{i,ideal} + \frac{V_{ref}}{2^n} * \sum_{k=1}^i \frac{\Delta R_k}{R}$$

Above Equation is not of much importance by itself, but it can be used to help determine the nonlinearity errors.

## II. DAC ARCHITECTURE

### 1. Introduction

The value of CMOS resistor can be defined from below equation with the equation. And the resistance of a MOSFET operated in the linear mode is given by

$$R = V_{DS} / I_D$$

Where the  $V_{DS}$  = drain-source voltage

$I_D$  = drain current

The drain-source voltage and the drain current are related by

$$I_D = \mu_n C_{ox} \frac{W}{L} * (V_{GS} - V_t) * V_{DS}$$

$$\text{For } V_{DS} \ll (V_{GS} - V_t)$$

Where value of  $W$  and  $L$  are variable, but their ratio will maintain the proper  $I_D$  value.

### 2. Operation of Designed DAC

An example of a 3-bit R-2R MOS converter is shown in figure A. in this figure R-2R cell shown with a possible combination with the bit current switch. All transistors in this system are equal.

Depending upon the input current  $2I$ , transistor M1 and M2 divide the input current  $2I$ .

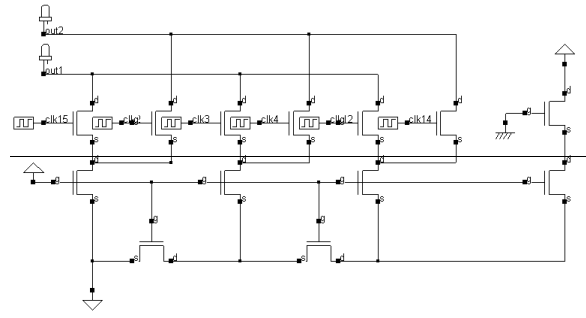


Fig A : 3 Bit R-2R MOS Converter

Transistors M1 and M2 can operate in saturated mode or in a triode mode. In saturated mode transistors M1 and M2 divide the input currents  $2I$  in to two equal currents  $I$ . in this case transistor M3 acts as a cascade transistor and supplies the output currents to the load.

At the moments transistors m1 and M2 are in triode region, then these transistors can be seen as a resistor with value  $R$ . in this case transistor M3 performs an equal resistor of value  $R$ . in this way R-2R network is implemented and with careful termination an accurate binary weighted current division is obtained. However it is possible to include the switches in to the network by adding transistor M4. at the moment data is high, then transistor M3 is used in the network as described before and the output current  $I_{out} = I_1$  is supplied to the load .at the moment data bar is high, then transistor M4 is used in the network as described before and the output current  $I_{out} = I_2$  is supplied to the load

By cascading method the basic elements of 3-bit converter can be designed as shown in fig A. the transistor system can be scaled depending upon the current value flowing through the individual stages. In this system trail current  $8I$  is divided by  $4I$ ,  $2I$ ,  $I$  and  $I$ . the extra current  $I$  is obtained in the last stage

Is supplied to the bias voltage and it is not required for the digital to analog converter. With large size of the division transistor it is possible to obtain 10 bit resolution with +/- 0.5% linearity.

An accurate switching of the current is required to obtain a small glitch when the digital to analog converter is switched around the MSB values .in the offset then transistor M4 is used in the network as described before and the output current  $I_{out} = I_2$  is supplied to the load

Biggest problem in R-2R DACs is mismatching of resistors values. This creates error in resolution. Resolution can be achieved by using operation amplifier or Low pass filter as a next stage. Proper switching is also very important.

III. OPERATIONAL AMPLIFIER ARCHITECTURE

The complete schematic of the ladder including the biasing for the op-amp is as per given below.

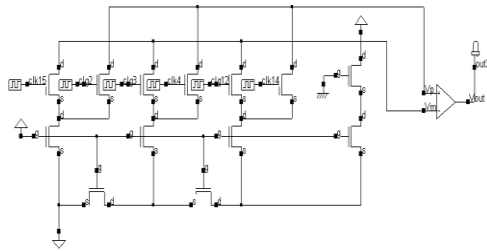


Fig B -3 bit R-2R ladder with operational amplifier

3.1 Open Loop Gain

The op-amp in this DAC is used in unity feedback configuration.

The closed loop gain for an op-amp is given by  $A/(A+1)$ , where A is the open loop gain

The error  $V_{out}/V_{in}$  is highest at the highest possible V which is  $V_{ref}=2V$

3.2 Gain Bandwidth

For a maximum operating frequency of 25 MHz, the output needs to be settled in 40ns

Again as in the calculation of the open-loop gain, the maximum speed is needed when the output voltage is the highest i.e. 5V

$$4.9976=5(1-e^{-(t/RC)})$$

$$RC=5.2ns$$

Or the gain bandwidth comes out to be 20 MHz

3.3 first stage of Operational amplifier

The first stage of the op-amp as shown in the fig gives a gain of about 45-60 from a common mode range of 3-4V.

Sizes of the transistors are given in fig.

The lengths of transistors had to be increased to provide high small signal output impedance for a high gain.

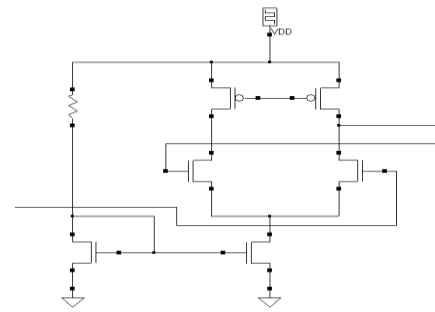


Fig C -1<sup>st</sup> stage of Op-Amp circuit

The second stage of the op-amp is acts as a gain stage and to provide enough current to drive a load of 20pF. The compensation capacitor was placed to improve the bandwidth of the amplifier.

Taking parasitic and load capacitance in to account, the maximum frequency of the second pole was calculated to be

$$1/RC = gm2/(C1+C2)$$

Where C1 and C2 are given from the figure as below

$$C1= Cgd1 + Cgd3 + Cdb1 + Cdb3 + Cgs5$$

$$C2= Cc+ Cdb5$$

$$C1= Cc+Cgd5$$

$$gm2= gm5$$

$c1+ c2$  can be taken as 20pf as parasitic are negligible as compared to load capacitor Hence  $gm2$  comes out to be 6.2 mA/V

And from the formula

$$gm = 2I/(Vgs -Vt)$$

I= bias current in second stage is calculated to be 2.3 mA

The 1st pole is arbitrarily chosen to lie at 16 KHz

And the 2nd pole is calculated to be 50 MHz

A resistor is placed in series with the capacitor so that the zero does not interfere in the transfer function of the gain.

The compensation capacitor is calculated to be 6.44 pF

And the zero resistors is calculated from the formula

$$Rz = (C1+Cc)/gm5*Cc$$



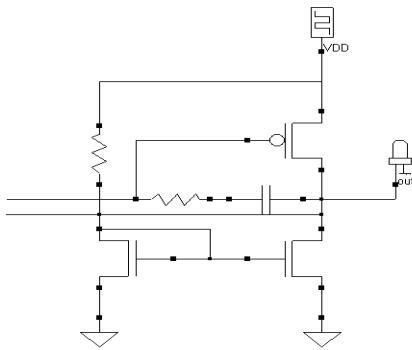


Fig D - 2<sup>nd</sup> stage of Op-Amp circuit

IV. PRE LAYOUT SIMULATION:

The characterization methods of the CMOS R-2R ladder D/A converter described in the previous section have been applied to a 3-bit D/A converter fabricated using the 0.13u CMOS process. The circuit diagram of the D/A converter is shown in Fig. A.

By applying input volts=2.0v as a pulse and digital input from 000 to 111 we achieved respective voltage and observe respective parameters like INL and DNL Offset, Glitch etc.

1. Integral Nonlinearity

INL = Output value for input code  $n$  - Output value of the reference line at that point.

Here input bits are from 000 to 111 which are represented here by 0 to 7. And Y axis shows ideal and actual INL. Maximum INL is 0.43. Offset error must be corrected otherwise it will be replicated in all stage as INL errors.

Value of resistor must be fixed so R and 2R will be generated. Main error is generated due to mismatch in the matching of resistors values.  $V_{DS}$  and  $I_D$  will be crucial parameter with W and L.

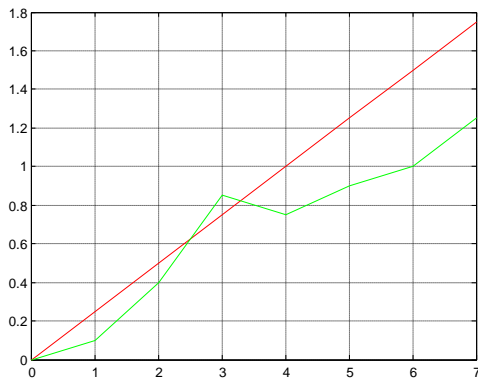


Fig E- Integration nonlinearity

2. Differential Nonlinearity

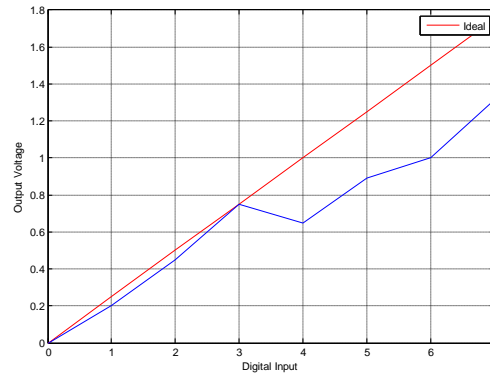


Fig F- Differential nonlinearity

Here input bits are from 000 to 111 which are represented here by 0 to 7. And Y axis shows ideal and actual DNL. Maximum DNL is 0.36.

4. Glitch

Here At the starting of input we find glitch which is biggest error and provide less resolution. This problem can be corrected by proper matching of resistor value and Width and length of the resistor.

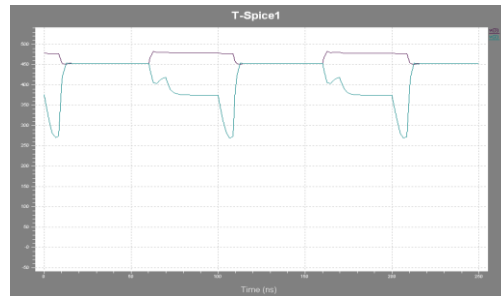


Fig G- Glitch at the output wave

5. Operational amplifier response

Here is the response of operational amplifier

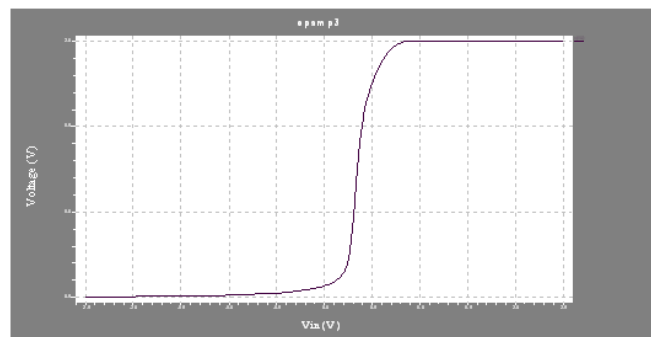


Fig I- Operational amplifier response

## V. Results and Conclusion

High resolution and linearity is also achieved with this all characteristics given below, this can be improved by adding more stages of R-2R.

TABLE-I  
Design Specification

Parameter	Design Specification & results
Technology	0.13um
Offset	32.0mV
INL	0.43
DNL	0.36
Glitch	Observed

The R-2R DAC is designed in TSMC 0.13um standard digital CMOS process and by using tanner tool.

## DAC DNL/INL summary

DAC choice of architecture has a significant impact on DNL. INL is independent of DAC architecture and requires element matching commensurate with overall DAC precision

And better INL and DNL can be achieved by proper matching of resistor s value in the CMOS circuit diagram.

## REFERENCES

- [1] R. Jacob Baker Harry W. Li David E. Boyce. CMOS Circuit Design, Layout and Simulation. IEEE Press Series on Microelectronics Systems, 2005.
- [2] Philip E. Allen and Douglas R. Hallberg. CMOS Analog Circuit Design. Oxford University Press, Inc USA-2002,pp.259-397, 2002.
- [3] Razavi, Behzad, "Data Conversion Sytem Design," IEEE Press, 1995, pp. 84 – 89.
- [4] Grebene, Alan B., "Bipolar and MOS Analog Integrated Circuit Design," John Wiley & Sons, 1984.
- [5] Johns, David A., and Martin, Ken, "Analog Integrated Circuit Design," John Wiley & Sons, 1997, pp. 471 - 473.
- [6] Hyung-rok Oh, et al., "Enhanced Write Performance of a 64-Mb Phase -Change Random Access Memory", *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 122-126, January 2006.
- [7] Kwang-Jin Lee, et al., "A 90nm 1.8V 512Mb Diode Switched PRAM with 266 MB/s Read Throughput", *IEEE Journal of Solid-State Circuits*, vol 43, pp. 150-162, January 2008.
- [8] Bedschi, Chiara Boffino, Edoardo Bonizzoni, Claudio Resta, Guido Torelli, "Staricase-down SET programming apporach for phase change memory" *Microelectronics Journal*, vol 38, pp. 1064-1069,2007.
- [9] Ferdinando 9. T.Nirschl, et al., " Write Strategies for 2 and 4 -bit Multilevel Phase-Change Memory", *IEEE International Electron Devices Meeting*, pp.461-464, December 2007.
- [10] D. Marche, Y. Savaria, and Y. Gagnon, "Compensated inverted R-2R ladder and compensation technique therefor," U.S. Patent Pending 11/411,110, May 19, 2006.
- [11] D. S. Karadimas, D. Mavridis, and K. A. Efstathiou, "A digitally calibrated R-2R ladder architecture for high performance digital-to-analog converters," in *Proc. ISCAS*, 2006, pp. 4779–4782.

- [12] D. Marche, Y. Savaria, and Y. Gagnon, "Laser fine-tuneable deep submicronCMOS 14 bit DAC," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 8, pp. 2157–2165, Sep. 2008.

# Study on effect of Calcium addition as a modifier in casting of LM25-SiC Composites

A. K Ganesa Balamurugan, B. Dr. K Mahadevan

**Abstract**—In this work, the effect of calcium addition as a modifier in casting of LM25-SiC composites was studied. The effect of Calcium addition was evaluated through microstructure and hardness of the cast composites and virgin LM25 alloy. LM25-SiC composites were prepared with three different weight percentage of Silicon Carbide (5%, 10% and 15%) using stir casting method. A constant amount (0.1% by weight) of calcium was added to the LM25-SiC composites and the virgin LM25 for comparison. The results reveal that the addition of calcium causes refinement of  $\alpha$ -aluminium dendrites in both LM25-SiC composites and virgin LM25 alloy with an appreciable improvement in the hardness due to grain refinement.

**Index Terms**—Calcium, LM24, Composites, SiC, modifier, Hardness

## I. INTRODUCTION

METAL matrix composites have market potential for various end applications, particularly in the automotive industry where the demand to use light weight materials has increased because of environmental issues [1]. Among the metal matrix composites, the most significant matrix materials are based on aluminium-silicon casting alloys [2]-[3]. The LM25 alloy (Al-Si1 Mg) is mainly preferred when there is a need for superior mechanical properties with ease in castability in order to achieve the desired shape and dimensions. The alloy is also used where resistance to corrosion is an important consideration [5]. However, the eutectic silicon in aluminium alloys is present as coarse polyhedral particles which accounts for inferior mechanical properties if left unaltered [6]. Hence, the modification of eutectic silicon by addition of an agent is of general interest in the manufacturing of Aluminium-Silicon castings to form fine eutectic silicon along with fine primary aluminium grains leading to improved mechanical properties [7]. Strontium and antimony are the commonly used agents for chemical modification of Aluminium-Silicon alloy castings [8]. Addition of small amounts of these elements to the melt changes the coarse and large needles of silicon into a fine and well rounded forms [9]. Similar to strontium, calcium also

belongs to the S-block group in the periodic table of elements. Hence in the present work it was decided to use calcium as an alternate which is cheaper and easily available material as the chemical agent for modification of Aluminium-Silicon alloy castings.

## II. EXPERIMENTAL PROCEDURE

The matrix material used in this work was LM25 (Al-Si7Mg) alloy. Green colour silicon carbide particles of size 250  $\mu\text{m}$  were used as the reinforcement material. The LM25-SiC composites with three different weight percentage (5%, 10% and 15%) of silicon carbide were casted using stir casting technique, the details given elsewhere [10]. A separate ingot of LM25 virgin alloy is prepared without reinforcement for comparison purpose. The above castings were made with calcium as the modifying agent. A constant 0.1% by weight of calcium in powder form was added to the melt with a short delay after the introduction of SiC particles. Totally eight castings; four modified and four unmodified were casted for comparison purpose. Specimens for microstructure analysis were cut longitudinally from the cast ingots. The specimens were first rough polished using successive grades of emery sheets and fine polished using sylvete cloth in a disc polishing machine. The samples so prepared were etched using Hydrofluoric solution and the microstructure images were taken using METSCOPE-1 metallurgical microscope. The hardness testing was carried out in Brinell hardness testing machine with a steel ball indenter of diameter 5mm and a load of 250 kgf. The hardness values were obtained by taking average of minimum of three indentations.

## III. RESULTS AND DISCUSSION

### A. Microstructure Analysis

The effect of calcium additions on grain size was investigate in virgin LM25 cast alloy and LM25-SiC composites of the three different weight percentage (5%, 10% and 15%). Fig.1a-Fig.1d shows the microstructure of unmodified virgin LM25 cast alloy and LM25-SiC composites. The dendrites in unmodified specimens are highly elongated. Fig.1e-Fig.1h shows the modified virgin LM25 cast alloy and LM25-SiC composites. The dendrites in

A. research scholar in the Department of Mechanical Engineering, Pondicherry Engineering College, Pondicherry-605014; e-mail: gbmpondy@gmail.com.

B. Professor in the Department of Mechanical Engineering, Pondicherry Engineering College, Pondicherry-605014. e-mail: mahadevan@pec.edu

modified specimens get refined and less elongated. This refining of dendrites was achieved by suppressing the growth of dendrites within the eutectic phase. The  $Al_2Ca$  precipitates initiate the nucleation of dendrites and  $Al_4Ca$  precipitates suppress the growth of dendrites [10]. Additionally, the pinning effect of Silicon Carbide particles was also suppressing the grain growth.

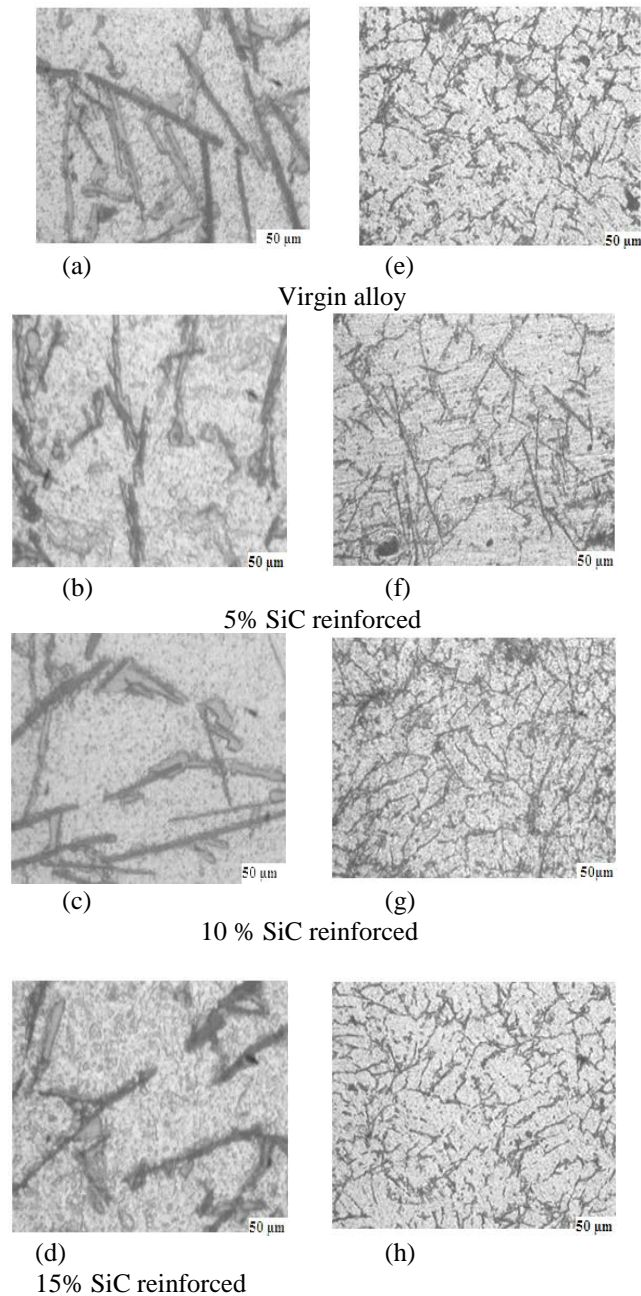


Fig.1 Micrographs of Unmodified (a-d) and Modified (e-h) composites and virgin alloys

**B. Hardness**

Fig.2 compares the hardness values of modified LM25-SiC and unmodified LM25-SiC composites. From the graph it is observed that increase in SiC particles content in LM25-SiC

composite increases the hardness of modified as well as unmodified LM25-SiC composites. Similarly, the modified LM25-SiC composites have comparatively higher hardness than the corresponding unmodified LM25-SiC composites, which is attributed to the grain refinement in modified samples LM25-SiC composites.

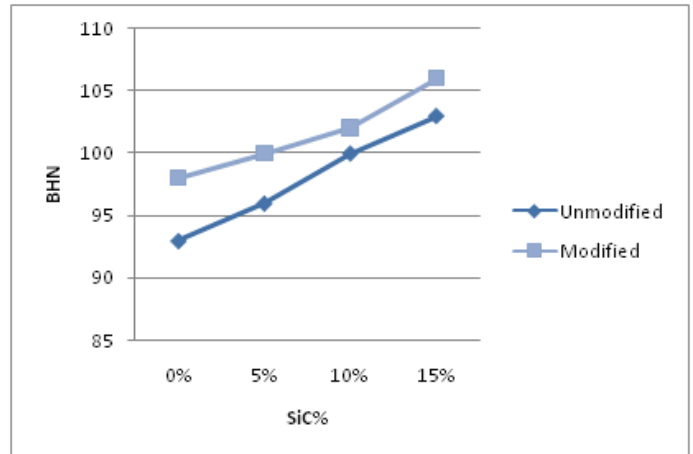


Fig.2 Brinell hardness values of unmodified and modified LM25-SiC composites and LM25 virgin cast alloys

**IV. CONCLUSION**

Addition of Calcium as modifier in LM25-SiC composites was studied. Due to addition of Calcium, refinement of grains occurred in composites and unreinforced alloy. The needles like silicon phase in the unmodified alloys were converted into fine particles by the addition of calcium. Increasing Silicon Carbide content in matrix increases the hardness of the composites and when compared to unmodified composites, modified shows the increase in hardness value.

**REFERENCES**

- [1] A.Ourdijini,K.C.Chew,B.T.Khoo,“Settling of silicon carbide particles in cast metal matrix composites”, Journal of Materials processing Technology, 116(2011), pg 72.
- [2] Jian Ku Shang, Weikang Yu, R.O.Ritchie, “Role of Silicon Carbide Particles in Fatigue Crack Growth in SiC-particulate-reinforced Aluminum Alloy Composite”, Material Science and Engineering A, 102(1988), pg 181.
- [3] T.J.A.Doel, M.H.Loretto, P.Bowen, “Mechanical properties of aluminium-based particulate metal-matrix composite”s, Composites. Number3. 1993. 24(1992), pg 269.
- [4] W.Zhou, Z.M.Xu, “Casting of Reinforced Metal Matrix Composites”, Journal of Materials Processing Technology, 63(1997), pg 358.
- [5] P. Mummery and B. Derby, “The influence of microstructure on the fracture behaviour of particulate metal matrix composites”, Materials Science and Engineering A, 135(1991), pg 221.
- [6] D.J.Lloy, H.Lagace, A.McLeod, P.L.Morris, “Microstructural Aspects of Aluminium-Silicon Carbide Particulate”, Material Science and Engineering A, A107 (1989), pp.73-80.
- [7] A.K. Dahle, K. Nogita, S.D. McDonald, C. Dinnis, L. Lu, Eutectic “modification and microstructure development in Al–Si Alloys”, Materials Science and Engineering: A, 413-414(2005), pg 243.

- [8] Hengcheng Liao , Min Zhang, Qichang Wu, Huipin Wang and Guoxiong Sun, “Refinement of eutectic grains by combined addition of strontium and boron in near-eutectic Al–Si alloys”, [Scripta Materialia](#), 57(2007), pg 1121.
- [9] Stuart D. McDonald, Kazuhiro Nogita and Arne K. Dahle, Eutectic grain size and strontium concentration in hypoeutectic aluminium–silicon alloys, [Journal of Alloys and Compounds](#), 422(2006), pg184.
- [10] D.P. Mondal, Nidhi Jha, Anshul Badkul, S. Das, M.S. Yadav, Prabhash Jain, Effect of calcium addition on the microstructure and compressive deformation behaviour of 7178 aluminium alloy, [Materials and Design](#), 32 (2011), pg 2803.

# 2D Fracture Analysis of an Excavator's Arm

A. Nitin Shukla, B. Prof. R R Trivedi, C. Prof. M. J. Mungla, D. Mr. Sridharababu I R

**Abstract**— Excavator's components are subjected to severe fatigue loadings and to calculate the fatigue life of welded regions in these components, the analysis is performed. Welds are most prone area for fatigue failure due to local stress concentration and geometrical changes. The fracture mechanics approach is used to calculate the life of the components once after the crack is established. The aim of the fracture mechanics simulation and calculations given here was the evaluation of the stress intensity factor and then by using compliance function, final crack length, at which the crack arrests is defined. The life is calculated afterwards. Hyper Mesh 8.0 is used as preprocessor and ABAQUS 6.7.3 with IDEAS 9 used as post processor.

**Key words:** Fatigue, stress concentration, fracture, stress intensity factor, compliance function.

## I. INTRODUCTION

Excavator's arm is one of the components of an excavator connected to boom from one end and with bucket to other end. The boom is attached with the help of two hydraulic cylinder one at the bucket cylinder bracket and other from the arm cylinder bracket. The bucket is attached to the arm from top boss and idler boss. All the plates of the arm welded to each other.

The basic function of the arm is to transfer the material from one location to destination with the help of bucket and boom. During the digging and dumping cycle the arm is subjected to fatigue loading. During the Digging-Dumping cycle the weld areas show a tendency of crack generation and propagation, which causes the failure.

It has been observed that due to this fatigue loading, the arm get fractured and then the failure occurs. To calculate its life after onset of the crack, the analysis has been performed. The stress analysis results were carried out as input and then various load calculations were done manually for crack analysis.

The paper describes two dimensional fracture mechanics approach to the excavator's arm at weld areas, thus a generalization of like cases is studied.

## II. MATHEMATICAL FORMULATIONS

### A. J-Integral in Two Dimensions

[1]The J-integral is usually used in rate-independent quasi static fracture analysis to characterize the energy release

associated with crack growth. The J-integral is defined in terms of the energy release rate associated with crack advance. In the context of quasi-static analysis, the J-integral is defined in two dimensions as

$$J = \lim_{\Gamma \rightarrow 0} \int_{\Gamma} n \cdot H \cdot q d\Gamma$$

where  $q$  is a unit vector in the virtual crack extension direction; and  $n$  is the outward normal to  $\Gamma$ .  $H$  is given by

$$H = WI - \sigma \cdot \frac{\partial u}{\partial x} \quad (2)$$

[1].  $H$  is a complex stress function where  $I$  denote the imaginary part of the function. For elastic material behavior  $W$  is the elastic strain energy; for elastic plastic material behavior  $W$  is defined as the elastic strain energy density plus the plastic dissipation, thus representing the strain energy in an "equivalent elastic material". This implies that the J-integral calculation is suitable only for monotonic loading of elastic plastic materials.

### B. Stress Intensity Factor Extraction

The stress intensity factors  $K_I$ ,  $K_{II}$ , and  $K_{III}$  play an important role in linear elastic fracture mechanics. They characterize the influence of load or deformation on the magnitude of the crack tip stress and strain fields and measure the propensity for crack propagation or the crack driving forces. Furthermore, the stress intensity can be related to the energy release rate (the J-integral) for a linear elastic material through

(3)

$$J = \frac{1}{8\pi} K^T \cdot B^{-1} \cdot K \quad \text{where } K = [K_I \ K_{II} \ K_{III}]^T \text{ and } B \text{ is called the prelogarithmic energy factor matrix. For homogeneous, isotropic materials } B \text{ is diagonal and the above equation simplifies to}$$

$$J = \frac{1}{E} (K_I^2 + K_{II}^2) + \frac{1}{2G} K_{III}^2 \quad (4)$$

where  $\tilde{E} = E$  for plane stress and  $G$  is the torsion modulus.

A.M.Tech- CAD/CAM –Institute of Technology, Nirma University, Ahmedabad

B. Associate Prof., –Institute of Technology, Nirma University, Ahmedabad

C. Asst. Prof., –Gandhinagar Institute of Technology, Gandhinagar

D. Project Manager, FEA Dept., L&T eEngineering Solutions, Vadodara

Table: 1 Geometrical specification of the section

	mm
Length of side plate	42
Width of plates	6
Weld dimensions	5×5

C. Compliance method [2]

Compliance method is followed to calculate the final crack length at the instability of crack by using weight function. The graph between weight function and crack-to-width ratio gives the value at which crack arrests and thus final crack length is calculated manually. The compliance method used the weight function Y as given below:

$$Y = K/\sigma \sqrt{(\pi \times a)} \quad (5)$$

Where  $\sigma$  is the value of applied stress, and ‘a’ is the value of crack length. Now by using the weight function and crack-to-width ratio, the maximum value of crack-to-width ratio is calculated at the verge of instability of the crack.

D. Paris-Erdogan Equation [2]

In quasi-static (LEFM) conditions, the most popular, empirical relationship between the crack-growth-increment per cycle (da/dN) and parameters of stress range and the instantaneous crack length is that proposed initially by Paris and Erdogan which gives:

$$da/dN = AK^m \quad (6)$$

Whose integration between initial crack length and final crack length gives the expression of life of the component.

III. FRACTURE MECHANICS APPROACH

A. Geometry and Model

The cross section of the excavator’s arm is extracted using the CATIA V5 software. The dimensions and 2D figure is given here in fig. 1. The side plate and bottom plate are meeting at this cross section and the area taken is where the maximum stresses found. Geometrical specification is listed in Table 1. A crack is embedded 180 degree opposite to the weld.

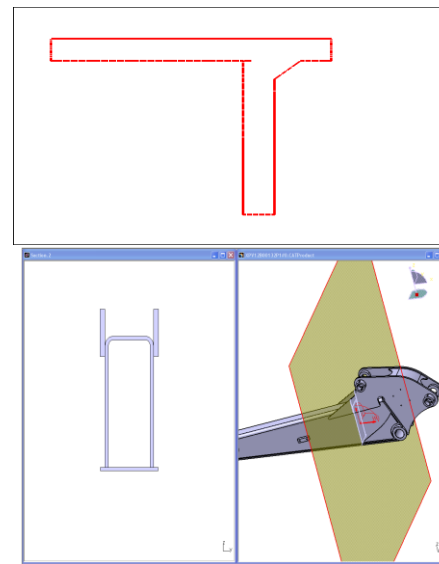


Figure 1 Section of the model

The crack is modeled with ring of fine quadrilateral elements of second order. A fine mesh for better results is created and total of five rings for five contour values has been made.

In a finite element model [3], each evaluation can be thought of as the virtual motion of a block of material surrounding the crack tip (in two dimensions). Each such block is defined by contours: each contour is a ring of elements completely surrounding the crack tip. These rings of elements are defined recursively to surround all previous contours. The rings of elements are 8-node biquadratic plane Stress quadrilaterals (CPS8R) as shown in Fig. 2. The element size is increased as you move away from the crack tip. The crack properties such as crack front and crack extension direction were defined.

In two dimensional fracture analyses, the crack front (first contour region) is the same as the crack tip. The crack extension direction was defined starting from the crack tip and going radially towards the external circle. Each contour provides an evaluation of the contour integral. The number of evaluations possible is the number of such rings of elements. 5 contours were specified to be used in calculating contour integrals.

B. Material Properties

The material is assumed to be linear elastic. Excavators operate at high loading conditions and considerable varying loads. They are fabricated various type of steels. For steel, the modulus of elasticity (E) of the ring material is  $2.1 \times 10^5$  Mpa.

C. Loading and Boundary Conditions

The loading condition for the static analysis and the crack analysis were tabulated in table 2, table3 and table4

respectively. The model is constrained symmetrically in x and y direction for the crack analysis and at the foot boss for static analysis to calculate the stresses.

Table 2 Load case 1 Digging

Component	Top Boss	Idler Boss	Bucket Cylinder Boss	Arm Cylinder Boss
Fx(in Kgf)	-560.63	-1036.36	1390.69	-9.68
Fy(in Kgf)	-769.89	352.67	276.11	-575.73

Table 3 Load Case 2 Dumping

Component	Top Boss	Idler Boss	Bucket Cylinder Boss	Arm Cylinder Boss
Fx(in Kgf)	-6076.43	214.23	5063.10	-6934.68
Fy(in Kgf)	341.12	822.26	352.52	7213.17

Table 4 Load for crack analysis

Tensile Stress	38.7165MPa
Tensile Load	232.299 N/mm
Section Modulus	6 mm <sup>3</sup>
Bending Load	464.598 MPa
Length of Side-Plate	42mm

D. The Final Model[4]

Except for the first five contours, the rest of the model has a coarse mesh; it was attempted to approximate the coarse meshes to square-shaped meshes. A part of the final meshed model is shown in Fig. 3. The final model was analyzed by ABAQUS/CAE 6.8-1 to give an estimate of J-integrals and stress intensity factors in the first 5 contours. After evaluation of stress intensity factor for various crack length, the weight function is calculated and tabulated in table 5. The graph between weight function and crack-to-

width ratio gives the maximum value of crack length at the verge of instability. The maxima function is applied to calculate the maximum value of crack length. This value of final crack length is used to calculate the remaining life of the component after onset of the crack.

Fig. 3 Meshed model

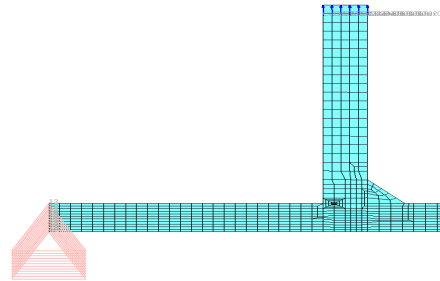
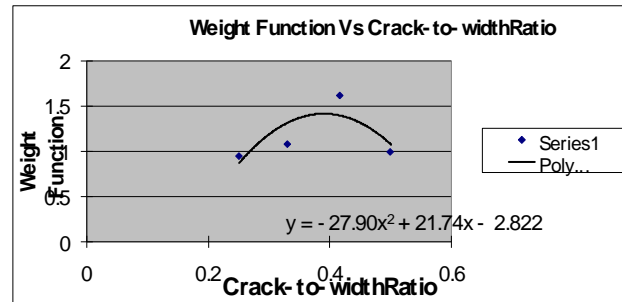


Fig 4 Weight function versus crack-to-width ratio



V. RESULTS AND DISCUSSION

The simulation led to the desired results. In linear elastic problems the first and second contours typically should be ignored as inaccurate, thus the average value of the J-integral is calculated based on contours 3, 4 and 5. The values of stress intensity factor were calculated since the crack front is very close to the symmetry axis, more refined meshes should be used to make the plane strain condition prevail locally around the crack front. The calculated values of the stress intensity factors *KI* and weight function is shown in Table 5 and the graph between weight function versus crack to width ratio is shown in figure 4. Using Paris-Erdogan equation the life is calculated as 3512 cycles. Later studies shall be developed on comparing simulation estimates with experimental results.



Crack length (a) (mm)	Crack-to-width Ratio (a/w)	K1 (Mpa*mm <sup>0.5</sup> )	S*(3.14*a) <sup>0.5</sup>	Weight Function (K1/S*(3.14*a) <sup>0.5</sup> )
1.5	0.25	160.03	168.09	0.952
2	0.33	210.4	194.09	1.084
2.5	0.4167	351.73	217	1.6208
3	0.5	236.63	237.7	0.998

## REFERENCES

- [1] C.H. Wang, Introduction to Fracture Mechanics, Aeronautical and Maritime Research Laboratory, DSTO-GD-0103.
- [2] J.F. Knott, P.A. Witney, Fracture Mechanics Worked Examples, The Institute of Materials, London, 1993
- [3] ABAQUS manual.
- [4] Razmi and Choupani, 2D Fracture Analysis of the First Compression Piston Ring, International Journal of Mechanical Systems Sciences and Engineering, 2007.

# Analytical and FEA Analysis for Existing Design of Top Frame of Web Guiding System

A. Jasmin. G. Patel, B. M. Y. Patil.

**Abstract**—This paper proposes a design study of web guiding system's frame structure. In frame structure mainly top frame of pivot type web guiding is checked for stress and deflection by analytically. Also it is difficult to create mathematical model for this top frame as structural members are loaded perpendicular to the frame plane. FE Analysis is done using SolidWorks simulation to find the Von-Mises stresses and deflection of top frame for a specific case study. Result are compared with analytical data. Deflection graphs are plotted to understand bending nature of structural members.

**Index Terms**—Web Guiding, FEA, Deflection, Von-Mises Stress.

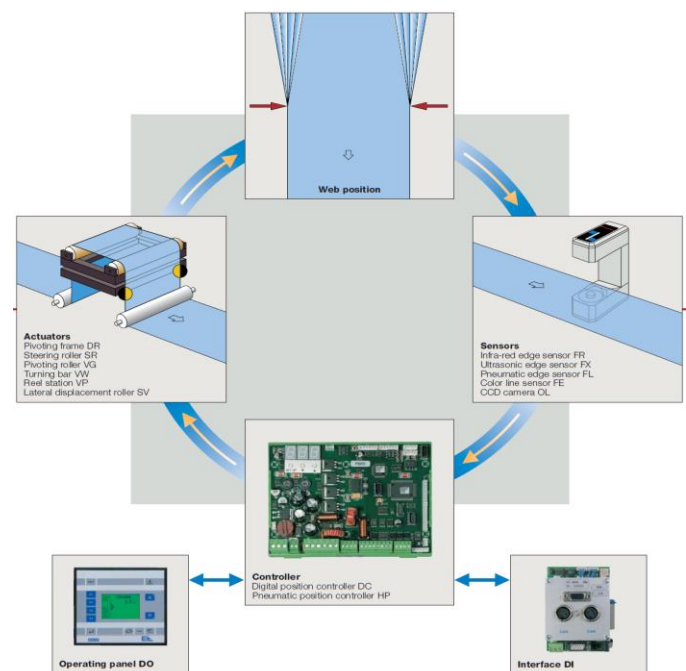
## I. INTRODUCTION

Materials such as paper, plastic film, foil and cloth often are produced in long, continuous sheets that are rolled up for more convenient handling and transportation. These rolls of material vary significantly in size and weight — ranging from 2 to 200 in. wide and weighing as much as several tons. The converting industry takes these continuous rolls of thin, flat materials — known as webs, threads them through processing machines (such as printing presses, laminating, coating and slitting machines) and converts or changes the web of material into an intermediate form or final product.

Web alignment is an important part of a converting operation as a moving web of material has a tendency to track off course and wander out of alignment during converting processes. To avoid these problems, engineers have developed a variety of automatic web-guiding systems that assure production accuracy and reduce waste. Web-guiding systems typically are positioned just before a critical stage on a converting machine (for example, just before a print station on a printing press).

Each type of web guiding system uses a sensor to monitor the web position for lateral tracking, and each has an actuator to shift the running web mechanically back on course whenever the sensor detects movement away from the set

path. Actuators may be pneumatic or hydraulic cylinders, or some kind of electromechanical device. These sensors may be pneumatic, photoelectric, ultrasonic, or infrared. The system's controls must put the output signals from the sensors in to a form that can drive the actuator.



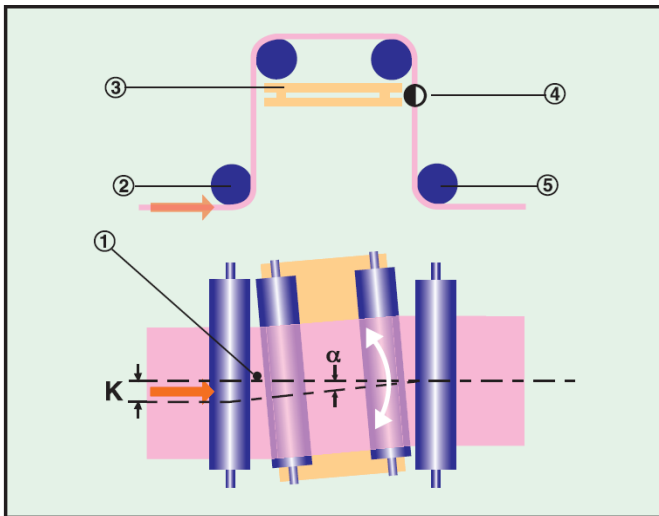
## II. PIVOTING FRAME SYSTEM

**A.Function:** The system is based on a pivoting frame with two path rollers. The imaginary pivoting point is located on the in-feed plane. Lateral web corrections may only be achieved by swiveling around this pivotal point. The prerequisite here is always sufficient tension for friction locking between the web and the guide roller.

**B.Application:** The greater the web tension, the elasticity module and the required correction, the longer the infeed, outfeed and transfer paths should be designed. Experience has shown that these paths should be the equivalent of 60-100% of the web width. The sensor should be located behind the guide roller as far as possible. Due to the short response times, increased guiding dynamics are achieved.

A. Post Graduate student in L.D. College of Engineering, Ahmadabad, Gujarat. (Corresponding Author, E-mail: jas\_mech84@yahoo.com)

B. AP, Department of mechanical engineering, L. D. College of Engineering, Ahmadabad, Gujarat. (Corresponding Author, E-mail: mypatil1@rediffmail.com)



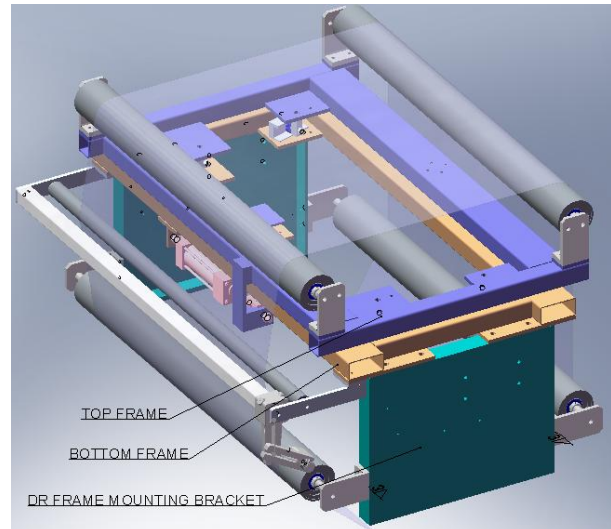
Web guiding frame structure can be divided in mainly three parts.

**C. Top Frame (Pivoting Frame):** Top frame has mainly two functions one is to support the rollers and another is to oscillate about the pivoting point to correct the web position. Size of top frame is customized and it is depends on roller face and correction span.

**D. Bottom Frame:** Bottom frame is base frame for top frame. This bottom frame is used to mount the supporting components for oscillating top frame. Size of bottom frame is customized and it is matched with customer’s machine frame. On the bottom frame there is provision to mount the sensors and servo centre.

**E. DR Frame Mounting Bracket:** Whole DR unit with the sub assemblies of top frame, bottom frame, entry roller and exit roller are mounted on DR frame bracket. Height of the bracket is depends on the max web width for correction. On this bracket necessary provision is given to level and to mount the DR unit on machine frame.

Below sketch shows the frame type structure for web guiding system.



**F. Design of Frame Structure**

Frame for guiding system are fabricated from the hollow rectangular or square sections or plate. Usually for below 1000 mm length plate type structure is preferable and above the 1000 mm length pipe type fabricated structure is used as it is more rigid.

Material of the structure is generally M.S. plate and steel hollow section (RHS/SHS) with YST24 grade. Fabrication of hollow section type structure required more quality concentration as all plates should be in plane and with 1mm / meter straightness and parallel to frame by 0.5 / 200 in X and Y direction. Traditionally 100 x 50 and 50 x 50 hollow section are used for web width ranging from 900 mm to 2000 mm.

The structure must have enough strength to with stand the roller weight and web tension. All the welded part must have good welding quality with proper care to avoid distortion in welding. All holes must bored or drilled with reference to the pivot point.

All structural members should be checked for the deflection and stress for its rigidity and strength. Factor of safety for the frame should be very conservative to avoid failure and better guiding. Weight of the frame structure will be consider during the selection of the actuator for the guiding. These both factors force for optimum structural frame design.

Traditional way of deciding frame sizes is described below as per E+L design standard. To calculate the top frame size, roller face and correction span are governing parameters.

$$\begin{aligned} \text{Top Frame Length (L}_{TF}) &= \text{Roller Face} + 70 \text{ mm} \\ \text{Top Frame Width (W}_{TF}) &= \text{Correction span} + 34 \text{ mm} \end{aligned}$$

These dimensions are based on consideration of using standard roller mounting brackets and with assumption that the roller shaft length is roller face plus 50 mm (design clearance).

To calculate the bottom frame and DR frame mounting

bracket are customize up to some extent and depends on the mounting position and space available. As per company standard bottom frame size is as per below.

$$\text{Min Bottom Frame Length (L}_{BF}) = \text{Top frame length} + 170 \text{ mm}$$

OR

$$\text{Machine Frame In To In} + 100 \text{ mm}$$

$$\text{Bottom Frame Width (W}_{BF}) = \text{Correction Span} - 86 \text{ mm}$$

III. CASE STUDY

Our main area of interest is frame structure of web guiding system mainly top frame because of other components such as roller design, brackets for roller mounting, sensor mounting brackets, hydraulic cylinder, power pack etc. are company standard. Also bottom frame is designed as per requirement of mounting on machine.

Our case study is on existing design having the following design input data.

A. Design Input:

- Web width Maximum : 900 mm
- Web width Minimum : 400 mm
- Lateral Movement of Web : ± 50 mm
- Web Tension : 40 Kg
- Material to Be Processed : Paper (40-80 gsm)
- Machine Frame In to In : 1130 mm
- Machine Frame Out to Out : 1250 mm

B. Top Frame Size as per Data

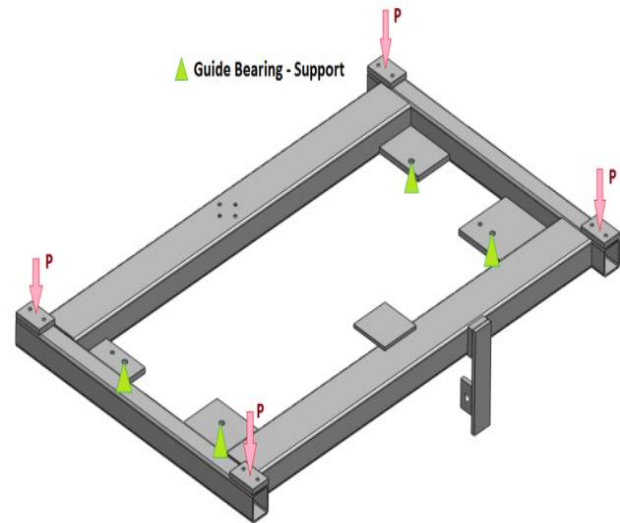
In exiting design the top frame is made up of standard hollow section of two sizes listed below. We will check this design for the deflection and the stress. Before that according to the standard parts available we fix the size (i.e. length and width) of the frame. Sections used for the structure are

- (i) 100 x 50 x 3 mm thick RHS
- (ii) 50 x 50 x 3 mm thick SHS

- a) Top Frame Length = Roller Face + 70 mm  
= 1050 + 70 mm  
= 1120 mm
- b) Top Frame Width = Correction Span + 34 mm  
= 726 + 34 mm  
= 760 mm
- c) On the top frame, roller mounting bracket and one end of the cylinder will be mounted and hence provision for mounting the cylinder and roller brackets should be there.

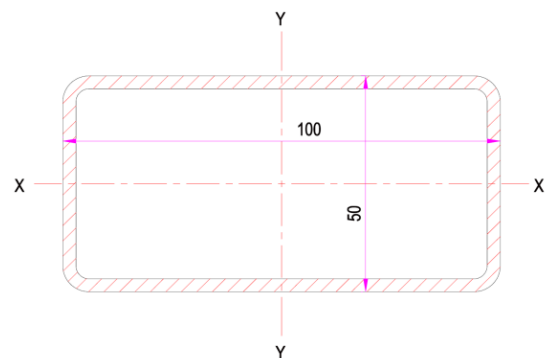
C. Design check for Stress and Deflection of Top Frame (Analytical)

We check the longitudinal member for bending stress and deflection. For this we take the case of both ends overhanging supports symmetrically with load of self weight and overhanging roller weight and tension load.



- To simplify the calculations we made some assumption.
- d) Beam is consider as simply supported at the guide bearings.
  - e) All guide bearings are equal distance from the centre line as there is small difference in distance compared to length.
  - f) Supports are assumed on the structural members for calculation.
  - g) We will check the stress and deflection for minimum distance between the supports as it is worst case.
  - h) Load due to cylinder force for oscillate the frame is neglected to simplify the calculations and also force value is small due to guide roller bearings.

As shown in drawing, the longitudinal section is 100x50x3 mm thick M.S. rectangular hollow section.



Section Properties are as follows.

Moment of Inertia

$$I_{xx} = 360571.36 \text{ mm}^4 \quad I_{yy} = 1064567.59 \text{ mm}^4$$

Section Modulus

$$Z_{xx} = 14422.85 \text{ mm}^3 \quad Z_{yy} = 21291.35 \text{ mm}^3$$

Material and its required physical properties are listed below (Reference from Tata Steel catalogue).

- i) Material = Yst 310 Grade Steel
- ii) Modulus of Elasticity,  $E = 2.060 \times 10^5 \text{ N/mm}^2$
- iii) Weight Density =  $7840 \text{ kg/m}^3$
- iv) Allowable Ultimate Strength,  $\sigma_u = 450 \text{ N/mm}^2$
- v) Allowable Yield Strength,  $\sigma_y = 310 \text{ N/mm}^2$

vi) Allowable Bending Stress,  $\sigma_b = 205 \text{ N/mm}^2$

There are two longitudinal structural members are used for the top frame and each contains two overhanging supports.

Total Frame Weight = 304.1 N

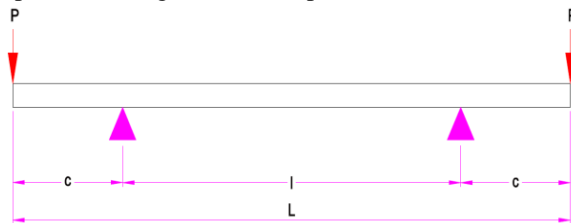
Weight of the Roller is combination of roller shell weight and shaft weight.

$$\begin{aligned} \text{i) Roller weight} &= (\text{Unit Shell weight} \times \text{Roller Face}) \\ &+ (\text{Unit Shaft Weight} \times \text{Shaft Length}) \\ &= (35.9 \times 1.050) + (38.5 \times 1.150) \\ &= 37.7 + 44.3 = 82.0 \text{ N} \sim \mathbf{90 \text{ N}} \end{aligned}$$

**Approx.**

$$\begin{aligned} \text{ii) Tension on Roller in worst case} &= \text{Unit Tension} \\ &\times \text{Web Width Maximum} \\ &= 436 \times 0.900 \\ &= 392.4 \text{ N} \end{aligned}$$

Each longitudinal structural member carry one roller weight, Tension in web in vertical direction and frame weight except its self weight on its end point.



$$\begin{aligned} \text{Load on each load point, } P &= (90 \div 2) + (392.4 \div 2) \\ &= 45 + 196.2 \\ &= 241.2 \text{ N} \sim \mathbf{250 \text{ N}} \end{aligned}$$

Stress at support and at all points between is critical and its value can be found from below equation.

$$\text{Bending Stress, } \sigma_{b1} = \frac{P c}{Z}$$

Where,

$\sigma_b$  = Bending Stress (N-mm)

P = Load at each load point (N) = 250 N

Z = Section Modulus = 14422.85 mm<sup>3</sup>

c = Distance between Load Point and adjacent support (mm) = 105 mm

$$\sigma_{b1} = 1.82 \text{ N/mm}^2$$

Deflection at load and at centre is critical and its value can be found as below.

*Deflection @ Load Point*

$$\delta_{LP1} = \frac{P c^2}{6EI} (2c + 3l)$$

$$I = L - 2c = 1070 - (2 \times 105) = 860 \text{ mm}$$

$$\delta_{LP1} = 0.0173 \text{ mm}$$

*Deflection @ Centre Point*

$$\delta_{C1} = - \frac{P c l^2}{8EI}$$

$$\delta_{C1} = 0.0326 \text{ mm}$$

Now, Stress and deflection due to its self weight are as per below.

Stress at support and at all points between is critical and its value can be found from below equation.

$$\text{Bending Stress, } \sigma_{b2} = \frac{W c^2}{2ZL}$$

Where, W = Total Load due to its self weight per structural member (N) = 71.7 N ~ 75 N

$$\sigma_{b2} = 0.027 \text{ N/mm}^2$$

Deflection at load and at centre is critical and its value can be found as below.

*Deflection @ Load Point*

$$\delta_{LP2} = \frac{W c}{24EIL} (3c^2(c + 2l) - l^3)$$

$$\delta_{LP2} = -2.37 \text{ E} -6 \text{ mm}$$

*Deflection @ Centre Point*

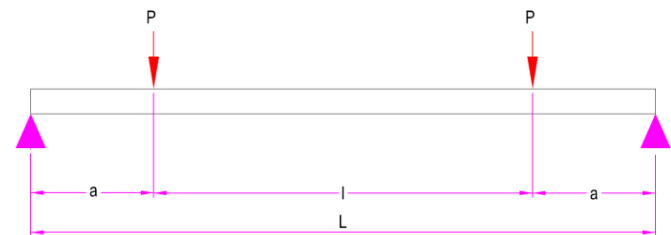
$$\delta_{C2} = \frac{W l^2}{384EIL} (5l^2 - 24c^2)$$

$$\delta_{C2} = 0.0062 \text{ mm}$$

Deflection and stress due to self weight are negligible compared to deflection and stress due to load. So now we consider only deflection and stress due to load for simplicity.

*For Inline Structural Member*

Similarly we can calculate the deflection and stress for the structural member in the direction of the web. Below data are derived from the excel program for this calculations.



Bending Stress,  $\sigma_b = 7.51 \text{ N/mm}^2$

Deflection @ Load Point = 0.078 mm

Deflection @ centre = 0.014 mm

IV. FE ANALYSIS REPORT OF TOP FRAME

A. Analysis type : Static

B. Mesh type : Solid Mesh

C. Material Properties:

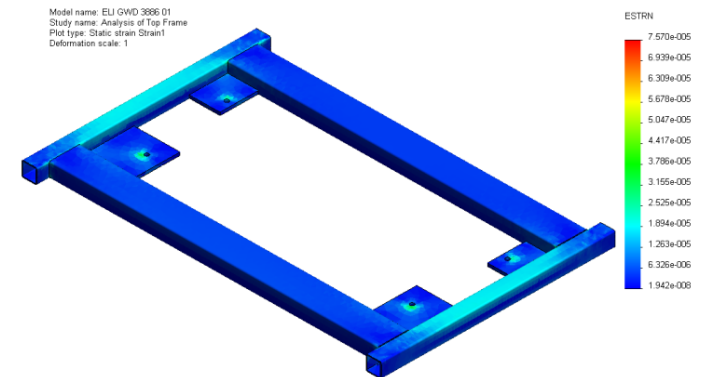
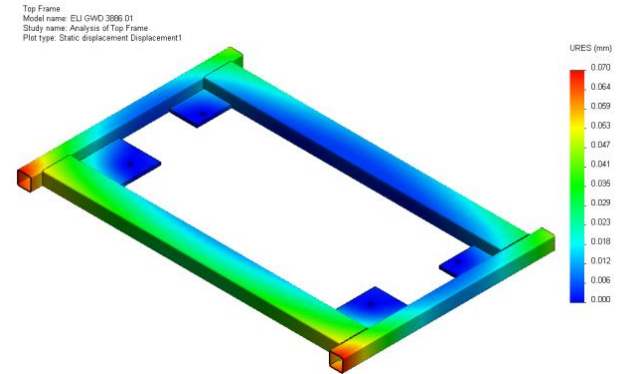
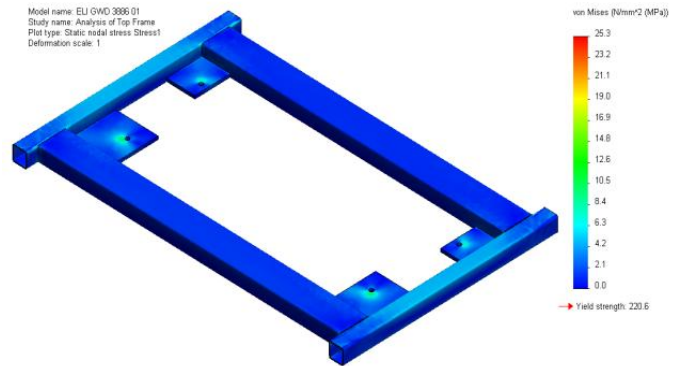
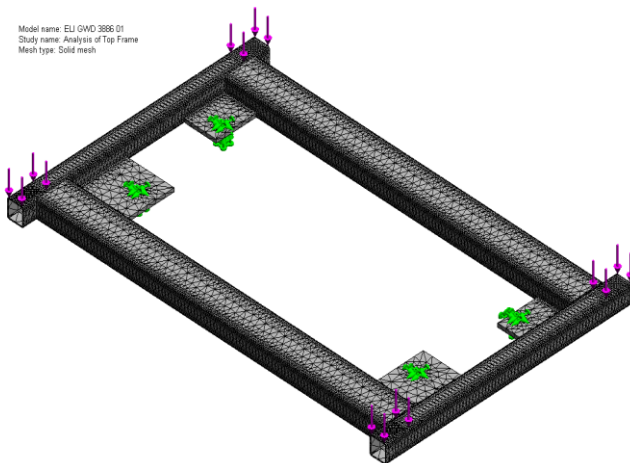
Material name : [SW]M.S. (Plain Carbon Steel)

Material Model Type: Linear Elastic Isotropic

Property Name	Value	Units
Elastic modulus	2.1e+011	N/m <sup>2</sup>
Poisson's ratio	0.28	NA
Shear modulus	7.9e+010	N/m <sup>2</sup>
Mass density	7800	kg/m <sup>3</sup>
Tensile strength	3.9983e+008	N/m <sup>2</sup>
Yield strength	2.2059e+008	N/m <sup>2</sup>

D. Loads and Restraints:

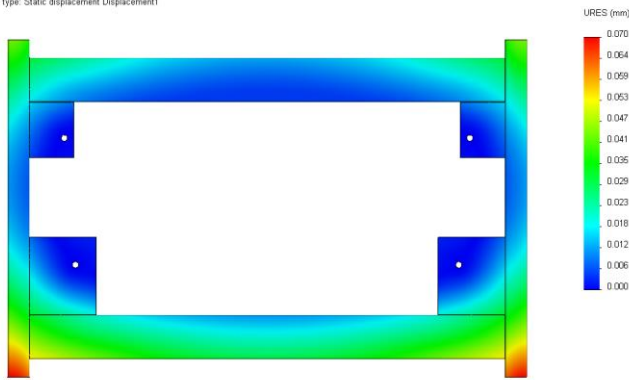
The top frame is restrained as a fixed at all four edges where the guide bearing is bolted as shown. 250 N force is applied on each of the four faces of the sections on which the roller mounting brackets are mounted.



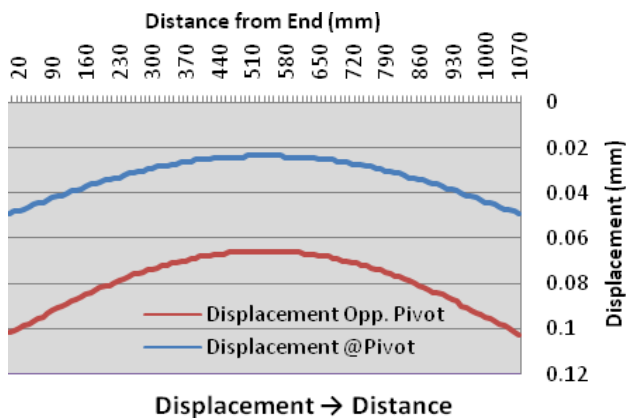
E. Study Results:

	Von-Mises Stress	Resultant Displacement	Equivalent Strain
<b>Min</b>	0 N/mm <sup>2</sup>	0 mm	1.94E-08
	Node: 57822	Node: 13	Element: 6464
<b>Max</b>	25.2967 N/mm <sup>2</sup>	0.0702315 mm	7.57E-05
	Node: 55895	Node: 71078	Element: 26376

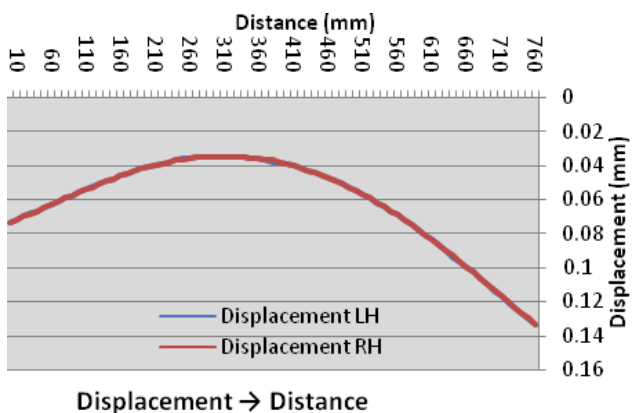
Top Frame  
 Model name: EJI\_GWD\_3886\_01  
 Study name: Analysis of Top Frame  
 Plot type: Static-displacement Displacement1



F. Study of displacement of Structural Member for Top Frame:



Graph shows the deflection of the longitudinal member at pivot point and the opposite to the pivot point. It is clear that the deflection tendency of both the members are same but the deflection value is different. This difference is due to the distance difference between the supports from the natural axis of member



In this graph the deflection of the inline member at both

side is shown. It is clear that the deflection tendency of both the members are same with the same deflection value.

G. Result and Discussion

	Result	Maximum Value
Analytical Values	Bending Stress	10.52 N/mm <sup>2</sup>
	Deflection	0.11 mm
Using FEA Software	Von-Mises Stress	25.29 N/mm <sup>2</sup>
	Resultant Displacement	0.07 mm

From the above result we can see that there is big difference in frames stress values but the deflection values are very near when comparing the analytical and FEA result. One of the reason of this difference is mathematical model of the top frame is not only the structural member but also there is plates are used to support the top frame and it is difficult to incorporate these plate in the calculation. For such analysis cases the FE analysis is useful.

VI CONCLUSION

Design of top frames is studied for deflection and stress to find the scope of the redesign and optimization of the design. Top frames is fabrication components and its cost is directly depends on their weight. Value of stress and deflection shows that the top frame can be re-design to reduce the weight of the frame to optimize it for cost.

REFERENCES

- [1] "Instruction Manual for Web Guiding" By – Erhardt + Leimer India Pvt. Ltd.
- [2] Design Standards for Web Guiding – By – Erhardt + Leimer India Pvt. Ltd.
- [3] "The Mechanics of Web Handling" By –David R. Roisum
- [4] "The Mechanics of Roller" By –David R. Roisum
- [5] "Machine Tool Design Handbook" By CMTI.
- [6] "Machine Design" By V. B. Bhandari.
- [7] "Machine Design an Integrated Approach" By Robert. L. Northon.
- [8] "Selecting a Web Guiding System" By – Bruce A. Feiertag
- [9] "Unwind and Rewind Guiding" By – Kenneth I. Hpcus

# Effect of Fuel Injection Advance on Pressure and Temperature of Four-Stroke Compression Ignition Engine by Phenomenological Modeling

A. Makwana V. M., B. Gandhi J. D., c. Modi A. J.

**Abstract**—This paper presents the effect of fuel injection advance on engine performance using a diesel engine running on decane fuel. In present work, a simulation has been done for four stroke compression ignition engine for progressive combustion to understand the effect of fuel injection on performance of four-stroke compression ignition stationary engine. Result of progressive combustion simulation was validated with result available in literature. Temperature (T) - crank angle ( $\theta$ ) diagrams plotting was done for progressive combustion for different angle of fuel injection. We observed that maximum temperature was 3190.13 K at crank angle in degree 366.092 when fuel is injected at an angle 333 degree crank angle for progressive combustion. We observed that maximum temperature was 3091.40 K, 2975.93 K, and 2824.49 K at crank angle in degree 368.64, 375.06, 403.02 respectively when fuel is injected at 340°, 350°, 360° crank angle respectively for progressive combustion. Result of progressive combustion simulation presented by plotting curve. Variations in pressure w.r.t. volume and crank angle are presented in graphical form.

**Key Words:**-Phenomenological Modeling, Progressive combustion, Advance Fuel Injection

## I. INTRODUCTION

IN engineering, computer simulation can give major contribution to engine design at different levels of general studies, or detail, corresponding to a different stages of model development [1]. As of now two basic types of model: Thermodynamic and Fluid Dynamic. Thermodynamic model on base of energy conservation are categorized as: zero-dimensional, phenomenological and quasi-dimensional. Phenomenological model based on each individual process occurring in engine cycle such as suction, compression, fuel injection, mixture formation, heat release, heat transfer, emission foramtion, expansion and exhaust. Fluid Dynamic model are often called as multi-dimensional; model due to their inherent ability to provide detailed geometrical information on the flow field, based on the solution of the governing flow equation.

The objective of the present work is to analyze the effect of

fuel injection timing on performance of single cylinder four stroke direct injection diesel engine using a computational thermodynamic model and progressive combustion. The CI engine is constant air supply engine. Fuel injection timing greatly affect on the combustion efficiency, ignition delay period hence on the efficecy and power. For simplitcy, phenomenological and single zone model has been choosen in this present work [2]. It is assumed that there is spatial uniformity of pressure, temperature and composition of the cylinder content at each crank angle. The analysis of model covers of compression, combustion and expansion processes. The analysis of the effect of timing of injection by covering the effects of heat losses, friction and temperature-dependent specific heats to obtain much closer approach to actual performance. This new developed model predicts in-cylinder temperatures and pressures as functions of the crank angle ( $\theta$ ).

The injection timing is very important for solid injection. In solid injection there is always time lag between the injection and time of combustion of injection fuel. This is inherent property of solid injection regardless of the shape and construction of nozzle. Solid injection generally gives coarse atomization and therefore requires earlier injection timing from 10 to 30° before TDC. The angle of injection advance in open combustion system varies from 10 to 30° and in precombustion chamber, it varies 30 to 45°. Fuel injection advance increasingly (relative to TDC) as the engine speed increases. In variable speed engines, provision made for automatic increase angle of advance with increasing speed. The correct timing advance for a given engine speed will allow for maximum cylinder pressure to be achieved at the correct crankshaft angular position.

## II. NOMENCLATURE

$A_c$	cylinder surface area, $m^2$
$A_h$	cylinder head area, $m^2$
$d_i$	bore of cylinder, m
$h_s$	intake specific enthalpy, J/kg
$h_d$	outlet specific enthalpy, J/kg
$K_a$	thermal conductivity of air, W/m K
$K_c$	thermal conductivity of cylinder material, W/m K
$l$	connecting rod length, m
$m_s$	intake mass flow, kg
$m_d$	outlet mass flow, kg
$m_f$	mass fuel, mol

A. working as Assistant Professor in Department of Mechanical Engineering in Government Engineering College, Bhuj-370001, Gujarat, India (\*corresponding author to provide; e-mail: makwana\_vinod@yahoo.com).

B. Assistant Professor in S.R. Rotary Institute of Chemical Technology (SRICT), Bh aruch-392001 (Gujarat – India ), (e-mail: jaiveshgandhi86@gmail.com).

C. Assistant Professor in Department of Mechanical Engineering in Government Engineering College, Bhuj-370001, Gujarat, India (e-mail: ashishjmodi@gmail.com).



n	no. of fraction fuel burned, mol
$n_p$	index of combustion
N	engine speed in revolution per minute
Pr	Prandtl number
P	pressure, N/m <sup>2</sup>
Q	heat transfer through cylinder walls & head, J
$R_c$	compression ratio
R	universal gas constant, J/mol K
Re	Reynold number
S	entropy, J/K
s	stroke length, m
$t_h$	thickness of cylinder head, m
$t_c$	thickness of cylinder, m
$T_{ca}$	air temperature at near cylinder head, K
$T_{cw}$	cooling water temperature, K
$T_{new}$	new temperature for next cycle, K
T	temperature, K
t	time, s
$U_c$	overall heat transfer coefficient of walls, J/m <sup>2</sup> K
$U_h$	overall heat transfer coefficient of head, J/m <sup>2</sup> K
U	total internal energy, J
u	specific internal energy, J/kg
V	volume, m <sup>3</sup>
$V_c$	cylinder volume, m <sup>3</sup>
w	angular speed, rad/sec
W	work done, J
$W_{net}$	net work done during cycle, J

*Greek Symbol*

$\Phi$	euivalence ratio
$\theta$	crank angle, ° CA
$\eta$	efficiency

*Subscripts*

a	air
c	compression
comb	combustion
disp	displacemnt
e	expansion
in	inflow
o	outside
1	before compression
2	after compression
3	end of combustion
4	end of expansion
5	end of exhaust

*Abbreviation*

A/F	air-fuel ratio
Act	actual
bth	brake thermal efficiency
BDC	bottom dead centre
CI	compression ignition
ith	indicated thermal efficiency
max	maximum
St	stoichiometric
TDC	top dead centre
VR	cut off ratio

III. MATHEMATICAL MODELING

As regards compression ignition engines, these models developed by applying the phenomenological and filling and emptying approach [3]. Consider the air and residual gas as control volume which is trapped between the cylinder and piston. If we neglect the potential and kinetic energy then, first law for control volume [4],[5].

$$\dot{Q}_{in} + \dot{W}_{in} = \frac{dU}{dt} + \dot{m}_d h_d - \dot{m}_s h_s \quad (1)$$

*A. Mathematical Formulation of Compression Process*

Assuming the cylinder walls are cooled by water and cylinder heads is cooled by air.

$$\dot{Q} = A_c U_c (T - T_{cw}) + A_h U_h (T - T_{ca}) \quad (2)$$

We compute the overall heat transfer co-efficient at cylinder walls by

$$\frac{1}{U_c} = \frac{1}{h_i} + \frac{\ln \frac{d_o}{d_i}}{2 \times \pi \times k_c \times l} + \frac{d_i}{h_o d_o} \quad (3)$$

Take  $d_o = d_i + 2t_c =$  outer diameter of cylinder,  $d_i =$  inner diameter of cylinder = d and  $h_i = h_c$ .

Equation (3) is

$$\frac{1}{U_c} = \frac{1}{h_c} + \frac{d}{2 \times k_c} \ln \left( 1 + \frac{2 \times t_c}{d} \right) + \left( \frac{d}{d + 2 \times t_c} \right) \frac{1}{h_o} \quad (4)$$

$h_o =$  heat transfer coefficient at outside (water side) approximately (150 W/m<sup>2</sup> K).

$h_c =$  inside heat transfer coefficient.

It can be determined by Mc Adam correlation for heating and cooling fluid in turbulent flow through pipes/tubes [6].

$$h_c = \frac{k_a}{d} \times 0.0245 \times Re^{0.8} \times Pr^{0.4} \quad (5)$$

Equation (4) gives the overall heat transfer coefficient at cylinder walls.

Cylinder surface area can be determined by

$$A_c = \frac{\pi \times d \times V_{tdc}}{A_h} + \left[ \frac{\pi \times d \times s}{2} \left( \frac{1 - \cos \theta}{2} + \frac{1}{s} \left( 1 - \sqrt{1 - \frac{s^2}{4 \times l^2} \times \sin^2 \theta} \right) \right) \right]$$

$$A_h = \frac{\pi \times d^2}{4} \quad (6)$$

$$(7)$$

To find overall heat transfer coefficient at cylinder head,

$$\sum Rt = \frac{1}{h_i A_i} + \frac{t_h}{k \times A} + \frac{1}{h_o A_o} \quad (8)$$

But  $A_i = A_o = A$ ; we get

$$\frac{1}{U_h} = \frac{1}{h_c} + \frac{t_h}{k_c} + \frac{1}{h_h} \quad (9)$$

$h_h =$  outside (air side) heat transfer coefficient of head (20 W/m<sup>2</sup> K).

Refer (4), (6), (7), (9), and (2), we can calculate heat transfer through cylinder.

Cylinder volume can be determined by

$$V_c = V_{tdc} + V_{disp} \left[ \frac{1 - \cos\theta}{2} + \frac{1}{s} \left( 1 - \sqrt{1 - \frac{s^2}{4l^2} \times \sin^2\theta} \right) \right] \tag{10}$$

$$\frac{dV_c}{d\theta} = \frac{V_{disp}}{2} \times \left[ \sin\theta + \frac{s \times \sin 2\theta}{4 \times l \times \sqrt{1 - \frac{s^2}{4l^2} \times (\sin\theta)^2}} \right] \tag{11}$$

$$\dot{W} = -P \frac{dV_c}{dt} = -P \times w \times \frac{dV_c}{d\theta} \tag{12}$$

$$\frac{dU}{dt} = \frac{d(mu)}{dx} = m \frac{du}{dt} + u \frac{dm}{dt} \tag{13}$$

But net mass flow rate in Control Volume

$$\dot{m}_s - \dot{m}_d = \frac{dm}{dt} \tag{14}$$

But during compression net mass flow rate is zero.

Equation (13) becomes

$$\frac{dU}{dt} = m \frac{du}{dt} \tag{15}$$

According to first Tds equation [3], we get

$$TdS = C_v dT + \left[ T \left( \frac{dP}{dT} \right)_v \right] dv \tag{16}$$

But for closed system,

$$TdS = du + PdV \tag{17}$$

Refer (16) and (17)

Equation (16) is

$$du + P \cdot dv = C_v dT + \left[ T \left( \frac{dP}{dT} \right)_v \right] dv \tag{18}$$

$$du = C_v dT + \left[ T \left( \frac{dP}{dT} \right)_v - P \right] dv \tag{19}$$

$$\text{But } \left[ \left( \frac{dP}{dT} \right)_v - P \right] = 0 \tag{20}$$

Equation (19) become

$$du = C_v dT \tag{21}$$

Refer (21) and (15),

Equation (13) is

$$\frac{dU}{dt} = m C_v \cdot \frac{dT}{dt} \tag{22}$$

Refer (1), first law applicable to cylinder during compression process

$$\dot{Q}_{in} + \dot{W}_{in} = \frac{dU}{dt} \tag{23}$$

Refer (22) and (23),

Equation (23) is

$$m C_v \cdot \frac{dT}{dt} = \dot{Q} - \dot{W} \tag{24}$$

$$w \frac{dT}{d\theta} = \frac{\dot{Q} - \dot{W}}{m C_v} \tag{25}$$

$$\frac{dT}{d\theta} = \frac{\dot{Q} - \dot{W}}{w \times m \times C_v} = \frac{\dot{Q} - \dot{W}}{M \times C_v \times n \times w} \tag{26}$$

$$\text{Where } \dot{W} = P \times w \times \frac{dV_c}{d\theta} \tag{27}$$

Hence Temperature is determined by solving equation numerically by using Euler's principle

$$T_{x+1} = T_x + \frac{dT}{d\theta} \times \text{stepsize} \tag{28}$$

Pressure can be calculated by equation

$$P_{x+1} = \frac{(N_a + N_x) \times R \times T_{x+1}}{V_c} \tag{29}$$

Where  $N_a$  = no. of moles of air

$N_x$  = No. of moles of exhaust residual gases

*B. Mathematical Formulation for Progressive Combustion Process*

In the present work, the combustion process is modeled using a single zone approach, which is based on a uniformly distributed heat releasing phenomenon. Spray combustion is not considered in detail. The combustion process as progressive combustion process. The following are species of interest during combustion : CO<sub>2</sub>, CO, O<sub>2</sub>, N<sub>2</sub>, H<sub>2</sub>O [7]. Start with putting the A/F, pressure and temperature at the end of compression and compute the adiabatic flame temperature T<sub>3</sub> [1] at constant volume by comparing the total enthalpy of reactant equal to total enthalpy of product. This comparison done by iteration method using Newton Raphson technique [8] and N<sub>p</sub> and N<sub>r</sub> found by the applying the chemical theory [1].

For the purpose of simulation, calculate pressure

$$P_3 = \frac{P_2 \times T_3 \times N_p}{N_r \times T_2} \tag{30}$$

P<sub>3</sub> = is the pressure that if all fuel burned instantaneously at tdc (i.e. at constant volume)

Where  $N_p$  = no. of moles of product  
 $N_r$  = No. of moles of reactant

$$V_3 = V_{tdc} \left( 1 + \frac{P_3' - P_2}{k \times P_2} \right) \quad (31)$$

Where  $k$  is the ratio of specific heats  
 $n$  = amount of fuel burned

Time rate of burning is given

$$\text{by, } \frac{dn}{dt} = \frac{dn}{dV} \frac{dV}{d\theta} \frac{d\theta}{dt} = \frac{1}{V_3 - V_{tdc}} V' \theta \frac{d\theta}{dt} \quad (32)$$

If the maximum burning rate is known or some hypothetical value is agreed upon, we can compute

$$\left( \frac{dn}{dt} \right)_{\theta=\theta_3} \text{ where } \theta_3 \text{ is the crank angle corresponding to } V_3$$

We compared it with the maximum allowable burning rate. If rate is not exceeding, then we can know the entire combustion process will proceed at the constant pressure. Hence the temperature at the end of combustion can be determined by

$$T_3 = \frac{P_2 \times V_3}{N_p \times R} \quad (33)$$

Refer (1), if the burning rate is more than maximum allowable burning rate then whole combustion process does not happened at constant pressure and it followed in step fashion. Consider combustion happened in DN steps (DN=constant taken between 30 to 60)

$$\Delta\theta = \frac{\theta_3 - 180}{DN}$$

(34) Using equation,

$$\Delta V = V' \theta \Delta\theta \quad (35)$$

And as long as the maximum burning rate is not exceeded,

$$\Delta n = \frac{\Delta V}{V_3 - V_{tdc}} \quad (36)$$

Change of pressure can be found during combustion by

$$\Delta P = \left[ P_{(\theta)} \times n_p \times \left( \frac{\Delta V}{V_c} \right) \right] + \left[ (P_3' - P_2) \times V_{tdc} \times \Delta n / V_c \right] \quad (37)$$

For all subsequent increment in time. The integration is completed when  $\sum \Delta n = 1$

The work of expansion during combustion is calculated by,

$$W_{comb} = \sum \left( P + \frac{\Delta P}{2} \right) \Delta V \quad (38)$$

### C. Mathematical Formulation for Expansion Process

Consider burnt gases - combustion product as control volume which is trapped between the cylinder and piston. All the steps are same in expansion process as compression process if we neglect the potential and kinetic energy . .Mathematical

modelling for the expansion process are same as compression process but have to be started from crank angle at which combustion ends.

Exhaust temperature can found by,

$$T_5 = T_4 \left( \frac{P_1}{P_4} \right)^{\frac{n-1}{n}} \quad (39)$$

Where  $n$  is index of expansion

Exhaust residual gas can be determined by,

$$N_x = \frac{P_1 \times V_{tdc}}{R \times T_5} \quad (40)$$

Amount of new intake fresh air found by,

$$N_a = \frac{P_1 \times V_{bdc}}{R \times T_1} - N_x \quad (41)$$

A new temperature for next cycle is can be determined by

$$T_{new} = \frac{R_c \times T_1}{R_c - 1 + \left( \frac{T_1}{T_5} \right)} \quad (42)$$

Engine specification are taken for the simulation are given below

- Bore (d) = 0.08 m and Stroke length (s) = 0.11 m
- Connecting Rod Length (l) = 0.23 m
- Compression Ratio,  $R_c$  = 16.5
- RPM (N) = 1500
- $T_a$  = 300 K and  $P_a$  = 1.01325 bar
- Fuel :  $C_{10}H_{22}$  and Molecular Weight of Fuel (WT)= 142
- No of Piston rings (NPR) = 3
- Piston Skirt Length (PSL) = 9.0 mm
- Inlet and Outlet Manifold Pressure = 1.01.325 bar
- Area of Exhaust valve (AEV) = 3.0  $cm^2$
- Area of Intake valve (AIV) = 2.5  $cm^2$

Engine specifications were taken for simulation from literature so we can validate the obtained result from the progressive combustion simulation with available result in literature [1].

For the purpose of simulation, the following expression has been chosen to represent following property [1].

$$h(T) = A + BT + C \ln(T) \quad (\text{kJ/mol}) \quad (43)$$

$$C_p(T) = B + C/T \quad (\text{kJ/mol K}) \quad (44)$$

$$C_v(T) = B - 8.314 + C/T \quad (\text{kJ/mol K}) \quad (45)$$

The coefficient A, B, and C taken for temperature range 400 to 1600 K is given in Table I [1], [2].

TABLE I

Gas	A	B	C	D
CO	299180.0	37.85	-4571.9	-31.10
CO <sub>2</sub>	56835.0	66.27	-11634.0	-200.00
H <sub>2</sub> O	88923.0	49.36	-7940.8	-117.00
N <sub>2</sub>	31317.0	37.46	-4559.3	-34.82
O <sub>2</sub>	43388.0	42.27	-6635.4	-55.15

The coefficient A, B, and C taken for temperature above 1600 K given in Table II [1], [2].

TABLE II

Gas	A	B	C	D
CO	309070.0	39.29	-6201.9	-42.77
CO <sub>2</sub>	93048.0	68.58	-16979.0	-220.40
H <sub>2</sub> O	154670.0	60.43	-19212.0	-204.60
N <sub>2</sub>	44639.0	39.32	-6753.4	-50.24
O <sub>2</sub>	127010.0	46.25	-18798.0	-92.15

IV. RESULT AND DISCUSSION

Assumption

1. The intake valve opens at TDC and exhaust valve open at BDC.
2. Neglecting gas exchange process.
3. Complete combustion of the fuel.
4. No physical delay, no dissociation during combustion, no time loss and no blow by loss.
5. Heat released from combustion is distributed evenly throughout the cylinder.
6. Enthalpy associated with pressure of injected fuel is usually not significant and hence ignored.
7. Spatially averaged, instantaneous (time resolved) heat transfer rates are used to estimate heat transfer to the cylinder walls.
8. The suction and exhaust process happened at constant pressure.
9. Air-fuel mixture) is considered an ideal gas.
10. No heat transfer and work transfer occurs between burned and unburned zones.

TABLE III

CODE VALIDATION AT  $\Phi=1$ ,  $R_c=16.5$  AND  $VR=3.75$

Variables	Result in Literature	Result by Present Code	Difference in %
P <sub>1</sub> [bar]	1	1.01325	1.31
V <sub>1</sub> [cm <sup>3</sup> ]	588.59	589	0.07
T <sub>1</sub> [K]	314.353	314.593	0.08
P <sub>2</sub> [bar]	50.34	51.12703	1.54
V <sub>2</sub> [cm <sup>3</sup> ]	356.72	360	0.91
T <sub>2</sub> [K]	959.139	962.053	0.30
P <sub>3</sub> [bar]	50.342	51.12703	1.54
V <sub>3</sub> [cm <sup>3</sup> ]	112.36	111	1.23
T <sub>3</sub> [K]	2572.085	2824.488	8.94
P <sub>4</sub> [bar]	6.365	6.33758	0.43
V <sub>4</sub> [cm <sup>3</sup> ]	588.59	589	0.07
T <sub>4</sub> [K]	1752.61	1855.797	5.56
P <sub>5</sub> [bar]	1	1.01325	1.31
V <sub>5</sub> [cm <sup>3</sup> ]	356.72	360	0.91
T <sub>5</sub> [K]	1216.472	1278.671	4.86
W <sub>net</sub> [J]	837.575	816.59	2.57
Indicated Power [ kW]	10.47	10.207	2.58
$\eta_{th}$ in %	43.943	46.373	5.24
Indicated MEP [bar]	14.95	14.769	1.23
Brake MEP [bar]	12.577	12.357	1.78
Brake Power [kw]	8.807	8.541	3.11
$\eta_{bth}$ in %	36.966	38.802	4.73

Table III shows that difference in power and efficiency between the result of literature and present code, is due to difference in inlet pressure condition and in combustion efficiency. In present work, combustion efficiency was taken 100% and in literature it was taken 80%. Differences in results between present work and literature can be ignored as differences are small.

TABLE IV

P, V, T AT  $\Phi=1$ ,  $R_c=16.5$  AND  $VR=3.75$

Variable	Fuel injected at 333° (CA)		Fuel injected at 340° (CA)		Fuel injected at 350° (CA)		Fuel injected at 360° (CA)	
	P (bar)	T (K)	P (bar)	T (K)	P (bar)	T (K)	P (bar)	T (K)
START OF COMPRESSION	1.013	314.397	1.013	314.774	1.013	314.733	1.013	314.593
END OF COMPRESSION	51.120	962.590	51.121	962.502	51.122	962.402	51.127	962.053
END OF COMBUSTION	53.909	2979.098	53.43219	2952.595	52.900	2923.033	51.127	2824.488
END OF EXPANSION	6.626	1962.438	6.565	1944.168	6.495	1923.785	6.267	1855.797
END OF EXHAUST	1.013	1340.432	1.01325	1329.892	1.013	1318.112	1.013	1278.671
MAXIMUM TEMPERATURE IN (K)	3190.13 AT 366.09 °CA		3091.40 AT 368.64 °CA		2975.93 AT 375.06 °CA		2824.49 AT 403.22 °CA	
MAXIMUM PRESSURE IN (bar)	57.728 AT 366.09 °CA		55.944 AT 368.636 °CA		53.857 AT 375.061 °CA		51.127 AT 403.22 °CA	

Table IV shows that temperature at end of combustion is increase as we increased fuel injection advance. This model on base of progressive comustion hence variation in pressure during combustion for diffrent angle of fuel injection is not marginalbe. Numerically maximum temperature and pressure is increase as we increased fuel injection advance but it's become later in terms of crank angle. If the injection advanced angle is increased, it increases the delay period because the pressure and temperature are lower when

injection begins. If the injection of fuel is too far advanced, the rate of pressure rise during auto-ignition is very high and cause detonation.

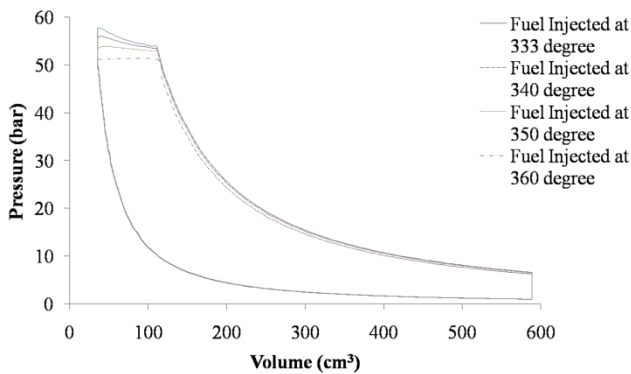


Fig. 1 P-V Diagram

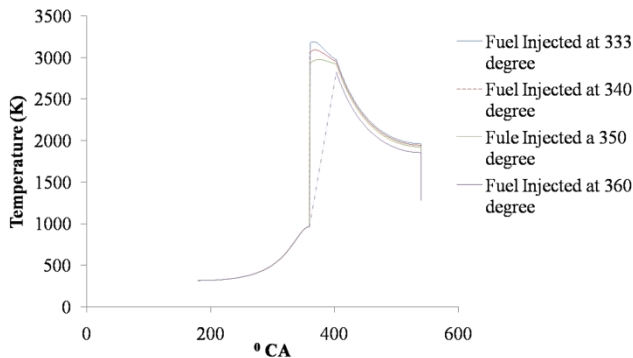


Fig. 2 T-θ Diagram

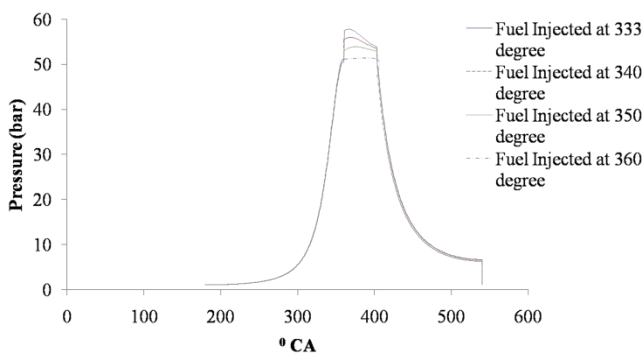


Fig. 3 P-θ Diagram

Fig. 1 shows the variation pressure with respect to volume. Fig. 2 and 3 show that variation of temperature and pressure with respect to crank angle for progressive combustion at  $\Phi=1$ ,  $N=1500$  and compression ratio=15 respectively. Fig. 2 shows that temperature instantaneously increase because we assume there is no ignition delay and it is homogenous and uniform distribution mixture. Fig. 2 and Fig. 3 show fuel injection is made advance, it increase the maximum temperature and increase the maximum pressure in cycle as we take adiabatic combustion. Pressure start falling after TDC due to movement of piston.

## V. CONCLUSION AND FUTURE WORK

1. When the fuel is injected at TDC, the pressure at the end of combustion is low due to expansion of gases.
2. Fuel injection is made advanced, then period of combustion in terms of crank angle is reduced.

3. Fuel injection is advanced by  $20^\circ$  and by  $10^\circ$  combustion take place at  $10^\circ$  and at TDC respectively, that generate moderate pressure.
4. Fuel injection is advanced by  $27^\circ$  and combustion take place at  $16^\circ$  before TDC and generate maximum pressure.

The degree of advance that any engine requires is dependent on a number of factors as compression ratio, A/F ratio, type of fuel, engine temperature and turbulence. The most important parameter which affect angle of advance is speed of engine. If the injection of fuel is too far advanced, the rate of pressure rise during auto-ignition is very high and causes detonation.

Marginal gap in result of efficiency between progressive combustion simulations and actual engine is due to valve operation, incomplete combustion, gas exchange process and temperature gradients in burnt zone and disproportionately high rate of heat transfer from the mixture that burns first during combustion. Progressive combustion simulation still popular due to their simplicity, low computational cost, reasonable accuracy and computationally undemanding requirements compared to more complex models [7-11].

Apart from the single-zone models, two-zone, and four-zone or even multi-zone models which furnish increased accuracy and flexibility for such complex phenomena as the formation of nitric oxide and soot in engine cylinders. The progressive combustion model do involve the fluid flow analysis and do not take care of premixed and diffusive combustion hence a detailed two zone model and three dimensional Computational Fluid Dynamics modeling would be essential to understand the phenomenon [12].

## VI. REFERENCES

- [1] Ganesan V, "Computer Simulation of Compression Ignition Engine Processes"; Tata McGraw-Hill, University Press (India) Ltd; New Delhi; 2000.
- [2] Makwana V.M., Lakdawala A.M., Shah P.P., "Numerical Determination of Delay Period of Four Stroke Engine Compression Ignition Engine by Zero Dimensional Modelling", Proceeding of TFMS-2012, Thermal Fluid and Manufacturing Science – 2012, Narosha Publication, New Delhi, ISBN 978-81-8487-202-6, p.p. 281-289.
- [3] Makwana V.M., Lakdawala A.M., Deshpande S. M., "Phenomenological Modelling of Four Stroke Compression Ignition Engine Processes with Dissociation Effect" Presented at 2<sup>nd</sup> International conference on Current Trends in Technology NUI-CONE-2011 08-10 December, organized by Institute of Technology, Nirma University, Ahmedabad, Gujarat, India, 2011
- [4] Nag P. K. , "Engineering Thermodynamics", Tata McGraw-Hill Company Ltd., Fourth Edition, 2008.
- [5] Deshmukh N.N., Malkhede D.N., "Experimental Performance Analysis of Single Cylinder Diesel Engine with blends of Food Grain Based and Non-Food Grain based Biodiesel", World Academy of Science, Engineering and Technology, 60, 2009
- [6] Kumar D. S. , "Heat and Mass Transfer", S. K. Kataria & Sons, Seventh Revised Edition, 2009.
- [7] Ramchandran Sundeep., "Rapid Thermodynamic Simulation Model of an Internal Combustion Engine on Alternate Fuels" Proceeding of the International MultiConference of Engineer and Computer Scientists 2009, Vol II, IMECS 2009, March 18-20, Hongkong, 2009
- [8] Rakopoulos C.D., Hountalas D.T., Rakopoulos D.C., Giakoumis E.G., "Experimental heat release rate analysis in both chambers of an indirect injection turbocharged diesel engine at various load and speed conditions." SAE paper no 2005-01-0926. Warrendale, PA: Society of Automotive Engineers Inc; 2005.

- [9] Ganesan V; "Internal Combustion Engine"; Second edition; Tata McGraw-Hill, New Delhi; 1999
- [10] Gatowski A., Balles E.N., Chun K.M., Nelson F.E., Ekchian J.A., and Heywood J.B. "Heat release analysis of engine pressure data." SAE Technical Paper 841359, 1984.
- [11] Ghojel Jami, Honnery Damon, "Heat Release Model for the combustion of diesel oil emulsion in DI diesel engines", Applied Thermal Engineering 25 , 2072-2085, 2005
- [12] Stone R. "Introduction to Internal Combustion Engines." 3rd ed. London: MacMillan; 1992.



**Makwana V. M** lives in bhuj and born in Gunjar, Gujarat, India on 29<sup>th</sup> October, 1977. He has completed gratuation in mechanical engineering from L.D. College of Engineering, Ahmedabad, Gujarat (India) in 1999 . Post graduate on thermal engineering from The M.S. University of Baroda, Vadodara, Gujarat (India) in 2011.

His area of interest includes Thermal Engineering, IC Engine and Computational Techniques in Mechanical Engineering. He

is currently works as an Assistant Professor at Government Engineering College, Bhuj (Kutch, Gujarat). He has more than 9 years of teaching and more than 3 years of industrial experience. He has published two technical paper and one international paper in journal/conference. Previous research is on fuel-air cycle simulation and current research is on preogressive combustion simulation.

Prof. Makwana is member of ISTE.

# Bond Energy Algorithm with Cross Weightage as decision parameter to form machine cell and part family

A. A M Gohil

**Abstract**—Formation of the machine cell and part family is the first step to be taken for implementation of cellular manufacturing concept. Cell formation is the first and most important step in successful implementation of cellular manufacturing. In this paper, Cross weightage is defined and proposed as the new decision parameter to place the rows and columns in the Bond Energy Algorithm of cell formation and to decide about the part families and machine cells. Solution generated by the proposed algorithm is compared with solutions available in the literature. The same algorithm is explained through example.

**Index Terms**— Bond Energy, Cross weightage, Exceptional cell, Machine Cell, Part Family.

## I. LITERATURE REVIEW

Group technology can be defined as a manufacturing philosophy identifying similar parts and grouping them together into families to take advantage of their similarities in manufacturing and design<sup>[1]</sup>.

Cellular manufacturing is the application of group technology concept in manufacturing. It is the physical division of the functional job shop's manufacturing machinery into Machine Cells. Machine cells processes a group of parts, known as Part Family, requiring similar machines, machining operations and toolings. Cellular manufacturing leads to reduced material handling, reduced tooling, reduced set-up time, reduced expediting, reduced in-process inventory, reduced lead time, improved human relations, improved operator expertise<sup>[2]</sup>.

The first step in cellular manufacturing is the formation of Machine Cell (MC) and Part Family (PF) which is known as Cell Formation. Cell formation approaches broadly can be classified as informal methods, part coding and analysis methods and production based methods. Several production based methods are developed and can be classed as cluster analysis, graph partitioning approaches, mathematical programming methods, heuristic and metaheuristic algorithms and artificial intelligence methods<sup>[3]</sup>.

McAuley (1972) used Single Linkage Cluster Analysis (SLCA) method using Jaccard similarity coefficient to solve the cell formation problem<sup>[4]</sup>. It was the first attempt to use similarity coefficient method to solve the cell formation problem.

J L Burbidge pioneered the concept of Production Flow

Analysis to solve cell formation problem. Production Flow Analysis is a method for identifying part families and associated machine grouping using the information contained on production route sheets rather than on part drawings<sup>[5]</sup>.

Kusiak developed Cluster Identification algorithm in which clusters are identified by crossing each entry by one horizontal and one vertical line<sup>[6]</sup>. Algorithm works well when there is no exceptional cell, but if the exceptional elements are present it fails to identify the cluster. Exceptional cell is the element belonging to the part which requires inter-cell movement for its manufacturing.

Boe and Cheng (1991) has proposed close neighbour algorithm for designing of cellular manufacturing systems in which the sum of the bond energy of rows and columns was used to make the decision about row and column placement<sup>[5]</sup>.

Miltenburg and Zhang has compared nine well-known algorithms which includes some of the well known algorithms like rank order clustering, similarity coefficient, modified similarity coefficient, seed clustering, and bond energy algorithm. Conclusion of the work was that no solution algorithm was found to be better than all other algorithms on the performance measures taken into consideration. Ideal Seed Non-Hierarchical Clustering (ISNC) algorithm proposed by Chandrasekharan and Rajagopalan (1986) was found to be better than others on smaller problems but not for the larger problems<sup>[8]</sup>.

McCormick, Schweitzer and White (1972) have developed an algorithm known as Bond Energy Algorithm (BEA)<sup>[9]</sup>. The algorithm works well for any size of the problems. However, initial row or column selection is done arbitrarily, which leads to a number of possible alternate solutions.

In this paper the new decision parameter termed as Cross Weightage is proposed to form part family and machine cell using a Bond Energy Algorithm. Algorithm works well on the problems of different difficulty level and leads to the unique and best possible solution.

## II. INTRODUCTION

Array based cluster analysis is one of the approach to solve the problem of Cell formation in which machine-part incidence matrix as shown in Fig.1 is used. It is based on the information available in the route sheet. Machines are

presented as rows and parts are presented as columns. It shows which machines are required to produce a particular part. An entry of "1" shows that the part requires processing on a machine and entry of "0" signifies that the part doesn't require processing on the machine.

The cell formation problem consists in arranging the rows and columns until the final solution matrix with entries of "1" arranged diagonally is obtained as shown in Fig. 2. Block diagonal structure in the solution matrix helps in identifying the Machine Cells and Part Families. Two clusters in the solution matrix indicates that two machines cells MC1 = {1, 3, 5, 6, 8} and MC2 = {2, 4, 7} can be formed. Also two part families formed are, PF1 = {2, 3, 5} and PF2 = {1, 4, 6}. Part family PF1 is processed on machine cell MC1 and part family PF2 is processed on machine cell MC2.

		Parts							
		1	2	3	4	5	6		
Machines	1	0	1	0	0	1	0	1	Machines
	2	1	1	0	0	0	0	2	
	3	0	1	1	0	1	0	3	
	4	1	0	0	1	0	1	4	
	5	0	1	0	0	1	0	5	
	6	0	1	1	0	1	0	6	
	7	1	1	0	1	0	1	7	
	8	0	1	1	0	1	0	8	
		1	2	3	4	5	6		
		Parts							

Fig. 1. Machine-part incidence matrix

		Parts							
		PF-1			PF-2				
		2	5	3	1	4	6		
Machines	MC-1	8	1	1	1				1
		3	1	1	1				2
		6	1	1	1				3
		5	1	1					4
		1	1	1					5
	MC-2	7	1			1	1	1	6
		4				1	1	1	7
		2	1			1			8
		1	2	3	4	5	6		
		Parts							

Fig. 2 Solution matrix

### III. BOND ENERGY ALGORITHM

#### 1) Bond Energy

Bond energy between any two columns  $j$  and  $j+1$  is given by the following formula:

$$BE_{ij} = \sum_{i=1}^m a_{ij} \times a_{i(j+1)}$$

(1)

Where,

$a_{ij}$  = machine-part incidence matrix element for the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column

$m$  = total number of machines or rows in the machine-part incidence matrix

Bond energy between two rows  $i$  and  $i+1$  is given by the following formula:

$$BE_{ij} = \sum_{j=1}^p a_{ij} \times a_{(i+1)j}$$

(2)

$p$  = is the total number of parts or columns in the machine-part incidence matrix

Rows and columns are rearranged till the matrix is arranged in such a way that total bond energy of the matrix is maximized. Total bond energy is given by the following formula:

$$BE_{total} = \sum_{i=1}^m \sum_{j=1}^p a_{ij} (a_{i(j-1)} + a_{i(j+1)} + a_{(i+1)j} + a_{(i-1)j})$$

(3)

#### 2) Bond energy algorithm

The bond energy algorithm works as follows:

1. Select column randomly from the machine-part matrix and place it in a new matrix.
2. Calculate the bond energy of each of the unselected columns with each of the selected columns. Select the column with maximum bond energy and place it in a new matrix as per the position determined while calculating the maximum bond energy. Repeat the process till all the columns have been placed.
3. Treat the matrix obtained in the step 2 as the source matrix and repeat the steps 1 & 2 for the rows representing machines.
4. After rearranging the rows the matrix obtained is the final solution matrix.

### IV. PROPOSED BOND ENERGY ALGORITHM WITH CROSS WEIGHTAGE AS DECISION PARAMETER

#### 1) Row or column weightage

Chan and Milner (1982) and King and Nakornchai (1982) have defined row weightage as follows:

$$w_{ri} = \sum_{j=1}^p a_{ij}$$

(4)

And column weightage is calculated as,



$$w_{cj} = \sum_{i=1}^m a_{ij}$$

(5)

Fig. 3 shows the weightage of each of the rows and columns. Weightage of the row 1 is calculated as,

$$w_{r1} = 0 + 1 + 0 + 0 + 1 + 0 = 2$$

Similarly the weightage of all other rows and columns can be calculated.

		Parts							
		1	2	3	4	5	6	$w_{ri}$	
Machines	1	0	1	0	0	1	0	2	1
	2	1	1	0	0	0	0	2	2
	3	0	1	1	0	1	0	3	3
	4	1	0	0	1	0	1	3	4
	5	0	1	0	0	1	0	2	5
	6	0	1	1	0	1	0	3	6
	7	1	1	0	1	0	1	4	7
	8	0	1	1	0	1	0	3	8
$w_{cj}$		3	7	3	2	5	2		
		1	2	3	4	5	6		
		Parts							

Fig. 3 Row and column weightage

2) Cross weightage

A new weightage factor defined and used for making the decision in BEA is “Cross weightage”. Cross weightage for the row is defined by the following formulae:

$$cw_{ri} = \sum_{j=1}^p a_{ij} \times w_{cj}$$

(6)

And cross weightage of column is defined as:

$$cw_{cj} = \sum_{i=1}^m a_{ij} \times w_{ri}$$

(7)

Fig. 4 shows the cross weightage for each of the rows and Fig. 5 shows the cross weightage for each of the columns. Cross weightage of the row 4 is calculated as,

$$cw_{r4} = 1 \times 3 + 1 \times 2 + 1 \times 2 = 7$$

Cross weightage of the column 3 is calculated as,

$$cw_{c3} = 1 \times 3 + 1 \times 3 + 1 \times 3 = 9$$

3) Proposed algorithm using cross weightage

1. Select the column with the maximum cross-weightage and place it first in a column sequence.
2. Calculate the bond energy of each of the unselected columns with the previously selected column. Select one of the unselected columns with maximum bond energy in the next stage.

- a. In case of the tie between two or more columns with same bond energy, select the column with the maximum cross weightage.
  - b. In case of the tie between two columns having same cross weightage, select the column with lower part or column number.
  - c. If decision is required to be made amongst the entries which are either zero or equal, look for the non-zero or higher bond energy entries in the unselected columns and place the same in the next stage.
  - d. Repeat the process till all the columns have been placed.
  - e. These steps will give the column or part sequence  $Seq_{column}$ .
3. Select the row with the maximum cross-weightage and place it first in a row sequence. Repeat the Step 2 to obtain the row or machine sequence  $Seq_{row}$ .
  4. Obtain the solution matrix by making the entries as per the column and row sequence obtained in the Steps 2 and 3 respectively.

		Parts									
		1	2	3	4	5	6	$w_{ri}$	$cw_{ri}$		
Machines	1	0	1	0	0	1	0	2	12	1	
	2	1	1	0	0	0	0	2	10	2	
	3	0	1	1	0	1	0	3	15	3	
	4	1	0	0	1	0	1	3	7	4	
	5	0	1	0	0	1	0	2	12	5	
	6	0	1	1	0	1	0	3	11	6	
	7	1	1	0	1	0	1	4	14	7	
	8	0	1	1	0	1	0	3	15	8	
$w_{cj}$		3	7	3	2	5	2				
$cw_{cj}$		9	19	9	7	13	7				
		1	2	3	4	5	6				
		Parts									

Fig. 4 Cross weightage of rows

		Parts									
		1	2	3	4	5	6	$w_{ri}$	$cw_{ri}$		
Machines	1	0	1	0	0	1	0	2	12	1	
	2	1	1	0	0	0	0	2	10	2	
	3	0	1	1	0	1	0	3	15	3	
	4	1	0	0	1	0	1	3	7	4	
	5	0	1	0	0	1	0	2	12	5	
	6	0	1	1	0	1	0	3	11	6	
	7	1	1	0	1	0	1	4	14	7	
	8	0	1	1	0	1	0	3	15	8	
$w_{cj}$		3	7	3	2	5	2				
$cw_{cj}$		9	19	9	7	13	7				
		1	2	3	4	5	6				
		Parts									

Fig. 5 Cross weightage of columns

4) Example to illustrate the working of proposed algorithm

- As shown in Fig. 5 the cross weightage of the column 2 is maximum (19) and hence it is selected first in a column sequence. Stage wise decision making is shown in Fig. 6.

Stage	Column	Columns					
		1	2	3	4	5	6
1	2	2	-	3	1	5	1
2	5	0	-	3	0	-	0
3	3	0	-	-	0	-	0
4	1	-	-	-	2	-	2
5	4	-	-	-	-	-	2
6	6	-	-	-	-	-	-

Fig. 6 Column placement

- As the bond energy of unselected column 5 is maximum with that of the selected column 2 it is selected next in the stage 2.
- In the stage 3 bond energy of unselected column 3 is maximum with the selected column 5, hence column 3 is selected in stage 3.
- Now the bond energy of the unselected column 1, 4 and 6 is zero hence there is a tie amongst them. But cross weightage of column 1 (9) is more than the cross weightage of columns 4 (7) & 6 (7) as shown in Fig. 5, hence column 1 is selected.
- Next bond energy of unselected column 4 and 6 is 2 with the previously selected column 1 and hence again there is a tie between them. Also the cross weightage of the column 4 and 6 is same i. e. 6. In such case column 4 is selected because column number  $j = 4$  is less than column no.  $j = 6$ .
- This procedure will give a column sequence as follows:

$$Seq_{column} = [2\ 5\ 3\ 1\ 4\ 6]$$

- Repeating the same process which is shown in Fig. 7, the row sequence obtained as follows:

$$Seq_{row} = [8\ 3\ 6\ 5\ 1\ 7\ 4\ 2]$$

Stage	Row	Rows							
		1	2	3	4	5	6	7	8
1	8	2	1	3	0	2	3	1	-
2	3	2	1	-	0	2	3	1	-
3	6	2	1	-	0	2	-	1	-
4	5	2	1	-	0	-	-	1	-
5	1	-	1	-	0	-	-	1	-
6	7	-	2	-	3	-	-	-	-
7	4	-	4	-	-	-	-	-	-
8	2	-	-	-	-	-	-	-	-

Fig. 7 Row placement

- Using the column and row sequence solution matrix obtained is as shown in Fig. 8. From the solution matrix it can be inferred that the two machine cells and two part families formed are,

$$MC1 = [1\ 3\ 5\ 6\ 8]$$

$$MC2 = [2\ 4\ 7]$$

$$PF1 = [2\ 3\ 5]$$

$$PF2 = [1\ 4\ 6]$$

		Parts							
		PF-1			PF-2				
		2	5	3	1	4	6		
Machines	MC-1	8	1	1	1				1
		3	1	1	1				2
		6	1	1	1				3
		5	1	1					4
		1	1	1					5
	MC-2	7	1			1	1	1	6
		4				1	1	1	7
		2	1			1			8
		1	2	3	4	5	6		
		Parts							

Fig. 8 Solution matrix

5) Test problems

To evaluate the proposed algorithm and to compare its performance with other cell formation methods, four data sets have been chosen from the literature as shown in Table I. The data set chosen includes part-machine incidence matrix of different difficulty level. It also includes exceptional element.

Solutions available in the literature and the proposed solution exactly match each other. One observed difference between the proposed solution matrix and the other solutions is that the solution matrix is not in block diagonal form; however the solution matrix available in the literature is in block diagonal form.

Table 4  
Cell formation problems

Sr. No.	Size	No. of Cells	Reference	Method	Solution Matrix
1	7×6	3	Kusiak <sup>[6]</sup>	Cluster Identification method	Fig. 9
2	12×10	3	McAuley <sup>[4]</sup>	SLCA	Fig. 10
3	8×20	3	Boe & Cheng <sup>[7]</sup>	Close neighbour	Fig. 11
4	7×9	2	Yasuda & Yin <sup>[10]</sup>	Dissimilarity measure	Fig. 12

		Parts									
		PF-1				PF-2			PF-3		
		2	3	5	8	1	6	4	7		
Machines	MC-1	1	1	1	1						
		5				1					
		7		1	1	1					
	MC-2	2					1	1			
		4						1			
	MC-3	3								1	1
		6									1

Fig. 9a Kusiak Solution matrix

		Parts									
		PF-1				PF2			PF-3		
		1	2	3	4	5	6	7	8	9	10
Machines	MC1	11	1	1			1	1			
		12	1	1			1	1			
		5					1	1	1		
		6					1	1	1		
	MC2	7				1	1	1			
		1	1	1	1	1					
		2	1	1	1	1					
		3	1	1	1						
	MC1	4	1	1							
		8								1	1
9									1	1	1
10									1	1	

Fig. 10b Proposed Solution matrix

Fig. 10 12x10 Machine-incidence matrix

		Parts									
		PF-1				PF-2			PF-3		
		3	5	8	2	4	7	6	1		
Machines	MC-1	7	1	1	1						
		1	1	1		1					
		5			1						
	MC-2	2							1	1	
		4								1	
	MC-3	3					1	1			
		6					1				

Fig. 9b Proposed Solution matrix

Fig. 9 7x6 Machine-incidence matrix

		Parts																				
		PF-1					PF-2					PF-3										
		3	4	6	7	18	20	2	8	9	11	13	14	16	17	19	1	5	10	12	15	
Machines	MC-1	7	1	1	1	1	1	1				1										1
		2	1	1	1	1	1	1					1									
		4	1	1	1	1	1	1														1
		8	1	1	1	1	1	1														
	MC-3	1	1						1	1	1	1	1	1	1	1	1					
		3	1						1	1	1	1	1	1	1	1	1					
5		1															1	1	1	1	1	
6			1														1	1	1	1	1	

Fig. 11a Boe & Cheng solution matrix

		Parts																				
		PF1						PF2						PF3								
		3	6	20	4	7	18	12	10	1	5	15	9	11	14	17	2	8	13	16	19	
Machines	MC-1	7	1	1	1	1	1	1	1			1										
		2	1	1	1	1	1	1							1							
		4	1	1	1	1	1	1				1										
		8	1	1	1	1	1	1														
	MC-3	1	1											1	1	1	1	1	1	1	1	1
		3												1	1	1	1	1	1	1	1	1
5		1															1					
6			1																			

Fig. 11b Proposed Solution matrix

Fig. 11 8x20 Machine-incidence matrix

		Parts										
		PF-1				PF2			PF-3			
		1	2	3	4	5	6	7	8	9	10	
Machines	MC1	1	1	1	1	1						
		2	1	1	1	1						
		3	1	1	1							
		4	1	1								
	MC2	5					1	1	1			
		6					1	1	1			
		7				1	1	1				
	MC3	8								1	1	1
		9								1	1	1
		10								1	1	
	MC1	11	1	1			1	1				
		12	1	1			1	1				

Fig. 10a McAuley Solution matrix

		Parts											
		PF-1				PF-2							
		1	3	9	5	6	8	2	4	7			
Machines	MC-1	2					1		1	1	1	2	
		4							1	1	1	4	
		6							1	1	1	1	6
		1				1	1	1					1
	MC-2	3	1	1	1	1							3
		7	1	1	1	1							7
5		1	1	1								5	
1	3	9	5		6	8	2	4	7				

Fig. 12a Yasuda and Yin solution matrix

		Parts									
		PF-1				PF-2					
		5	3	9	1	2	4	7	6	8	
Machines	MC-1	3	1	1	1	1					3
		7	1	1	1	1					7
		5	1	1	1						5
	MC-2	1	1						1	1	1
		2					1	1	1	1	2
		6					1	1	1		6
		4					1	1	1		4
		5	3	9	1	2	4	7	6	8	

Fig. 12a Proposed Solution matrix

Fig. 12 7x9 Machine-incidence matrix

[5] J. L. Burbidge "Production flow analysis for planning group technology"*Journal of operations management* , Vol. 10, pp. 5-27, 1991

[6] A Kusiak and W S Chow "Efficient solving of the group technology problem"*Journal of manufacturing systems*, Vol 6 No. 2, pp. 117-124, 1987

[7] W J Boe and C H Cheng "A close neighbour algorithm for designing cellular manufacturing systems" *international journal of production research*, vol. 29, No. 10, pp. 2097-2116 1991

[8] J Miltenburg "A Comparative evaluation of nine well-known algorithms for solving the cell formation problem in group technology" *journal of operations management*, vol. 10, No. 1, pp. 44-72 January 1991

[9] W T McCormick, P J Schweitzer and T W White "Problem decomposition and data reorganization by a clustering technique"*Operations Research*, vol. 20, pp. 993-1009, 1972

[10] K Yasuda and Y Yin "A dissimilarity measure for solving the cell formation problem in cellular manufacturing" *Computers and industrial engineering*, vol. 39, pp. 1-17, 2001

V. CONCLUSION

Following can be concluded out of the work presented in this paper:

1. The proposed algorithm gives a unique solution to a cluster formation problem as against multiple solutions possible using conventional BEA algorithm in which rows and columns are selected based on intuition.
2. The solution obtained is the best possible solution which gives comparable results with the solution obtained by other methods available in the literature.
3. Computer implementation of the algorithm is easier and requires only integer calculations as against floating point calculations required in other methods like similarity coefficient.
4. To obtain the machine cells and part families the solution matrix need not be in block diagonal form as preached in the literature.

VI. FUTURE WORK

Comparative study of this algorithm can be made with other algorithms on various grouping measures like total bond energy, number of exceptional elements, grouping efficiency, grouping efficacy, cell density etc.

VII. REFERENCES

[1] H. M. Selim, R. G. Askin and A. J. Vakharia, "Cell formation in group technology: Review, evaluation and directions for future research," *Computers industrial engineering*, vol. 34, No. 1, pp. 3-20, 1998.

[2] T J Greene and Randall P Sadowski "A review of cellular manufacturing assumptions, advantages and design techniques" *journal of operations management*, vol. 4, No. 2, pp. 85-97, February 1984

[3] G Papaioannou and J M Wilson "The evolution of cell formation problem methodologies based on recent studies (1997-2008): Review and directions for future research"*European journal of Operations Research*, 206, pp. 509-521, 2010

[4] J McAulley "Machine grouping for efficient production" *The production engineer*, pp. 53-57, February 1972

# Safety Considerations in Industrial Boilers

A. Chintan T. Barelwala,

**Abstract**— Technological development is taking place at a very fast rate in all the fields like mechanical, metallurgical, chemical, electrical and civil. The importance of industrial safety is realized because every year millions of industrial accidents. It is essential to follow some safety rules for the boiler operators. Hazards in boiler industries safety should be identified and safety devices should be provided to prevent dangerous boiler operating conditions from turning into disasters, only proper precautions and maintenance can prevent dangerous boiler operating conditions from occurring.

**Index Terms**— Industrial Safety, Hazards, Boiler, Accidents, Safety Valve, Personal Protecting Equipments, risk, safety analysis.

## I. INTRODUCTION

These days every man is surrounded by automobiles, explosives, noise and pollution etc. which may cause accidents. The importance of industrial safety was realized because every year millions of industrial accidents occur which result in either death or in temporary or permanent disablement of the employees and involve large amount of loss resulting from damage to property and wasted man and machine hours. The companies have to pay compensation to the injured workers during accidents.

Industrial safety is mainly concerned with minimizing hazards in the industries. Hazard is a state, physical or chemical having potential to injure the person or impairment of health. Industrial safety training helps workers on job site recognize hazards and take positive action to avoid them. In some cases, industrial safety education is mandatory, and every worker at a given job site must complete training to continue working.

## II. GENERAL AREAS OF SAFETY

These areas are common to all type of industries and they are; safety, accidents, dangerous occurrences, accident prevention, plant layout, house-keeping, material handling, lighting, storage, ventilation, ergonomics, slips, trips and

falls, fire safety, machinery safety, electrical safety, safety at work at heights, office safety, safety in the use of hand tools, food hygiene, personal hygiene, occupational health practice, occupational stress, occupational hygiene, first-aid, legal aspects of safety, duties of workers, role of supervisors on safety, personal protective equipments, safety helmet, safety policy, safety audit, safety education, safety committee and, safety and productivity.

## III. BOILER

A boiler is a device used to create steam by applying heat energy to water. Technically, Boiler means any closed vessel exceeding 22.75 litres in capacity which is used expressly for generating steam under pressure and includes any mounting or other fitting attached to such vessel which is wholly or partly under pressure when steam is shut off.

All types of boilers that are used in commercial and public facilities that produce steam (low or high pressure). A boiler is used wherever a source of steam is required. The form and size depends on the application: steam engines, industrial installations and power stations. But industrial installations and power stations will usually have a larger separate steam generating facility connected to the point-of-use by piping.



Figure 1 Industrial Boiler for supplying steam.

### A. EFFECTS AND CAUSES OF BOILER ACCIDENTS

Boiler systems are designed for safety and efficiency. The boiler operator is the key to safe boiler operations. Having knowledge about boiler systems and maintenance can ensure years of safe and reliable service. Accident is caused due to an explosion of a boiler or steam pipe or any damage to a boiler or steam pipe which is calculated to weaken the strength thereof so as to render it liable to explode.

Historically, boilers were a source of many serious injuries and property destruction due to poorly understood engineering principles. Thin and brittle metal shells can rupture, while poorly welded or riveted seams could open up, leading to a violent eruption of the pressurized steam. When water is converted to steam it expands in volume over 1,000 times and travels a down a steam pipes at over 100 km/hr. Because of this steam is a great way of moving energy and heat around a site from a central boiler house to where it is needed, but without the right boiler feed water treatment, a steam-raising plant will suffer from scale formation and corrosion.

At best, this increases energy costs and can lead to poor quality steam, reduced efficiency, shorter plant life and an operation which is unreliable. At worst, it can lead to catastrophic failure and loss of life. Collapsed or dislodged boiler tubes could also spray scalding-hot steam and smoke out of the air intake and firing chute, injuring the firemen who loaded coal into the fire chamber. Extremely large boilers providing hundreds of horsepower to operate factories could demolish entire buildings.

From the History, it shown that without proper operation and maintenance, boiler conditions and safety deteriorate causing potential hazards due to neglect and misunderstanding. Routine maintenance is well within the ability of most boiler operators. Some of the areas where trained professionals needed are:

- Leaking safety relief valves
- Feed water to boiler
- Steam leaks (steam systems)
- High stack temperatures (excess of 350°F)
- Insufficient heat for building
- Condensate dripping down stack or out the front of the boiler
- Constantly resetting of controllers and safety devices

Boiler accidents can occur when the boiler is allowed to operate without adequate water in the boiler. Proper functioning low water cutoffs are essential to prevent these types of accidents. Boiler damage can run from severe

buckling and deforming of the boiler to complete meltdown or potential boiler explosions.

These accidents occur when the boiler can no longer contain the excessive pressure allowed to build in the boiler. Excessive pressure accidents, even in small boilers, have been known to completely destroy a building. Safety measures required are: Rupture disc in addition to a safety valve so that when safety valve does not operate (mostly it happens so because of inadequate size and poor maintenance) the rupture disc will release the pressure. High pressure alarm and automatic depressurizing. Interlocking of the manhole cover with the manual depressure valve so that unless the pressure is fully discharged by venting, the manhole cover cannot be opened.

Fuel related accidents usually occur when there is a failure to purge combustible gases from the firebox before ignition is attempted. Leaking fuel valves can also be the cause of these accidents. If the operator notices any gas odor, the boiler should be shut down and the fuel supplier notified immediately.

#### IV. KEY BOILER-SAFETY FEATURES

Boilers have a variety of features designed to prevent accidents and keep them functioning at optimal efficiency:

##### *A. Safety valves.*

Safety valves are the primary safety feature on a boiler. Safety valves are designed to relieve all of the pressure generated within a boiler if other systems fail. Every steam and hot-water heating boiler must have at least one safety or safety relief valve of sufficient relieving capacity to meet or exceed maximum burner output.



Figure 2 Boiler Safety Valve

The ability of a safety valve to perform its intended function can be affected by several factors, including internal corrosion and restricted flow.

• **Water-level control and low-water fuel cutoff.** These two devices perform two separate functions, but sometimes are combined into one unit. It is important to ensure piping is open and free of scale or sludge buildup at all times. Cross tees allow piping to be cleaned and inspected easily. Low-water fuel cutoffs should be checked periodically for proper operation. Because this requires boiler water to be lowered to the minimum safe operating level, extreme caution should be used.

In addition to periodic tests of a low-water device, the float chamber on a water-level control and/or a low-water fuel cutoff should be flushed thoroughly to remove accumulated sediment. At least once a year, water-level controls and low-water fuel-cutoff devices should be disassembled, cleaned, and checked.

#### B. Water gauge glass.

A water gauge glass enables an operator to observe and verify the actual level of water in a steam boiler. If not properly cleaned and maintained, a gauge glass can appear to show a sufficient level of water when a boiler actually is operating in a low-water condition. A stain or coating sometimes develops on the inside of a gauge glass, where the gauge glass is in contact with boiling water. This stain can give the appearance of water in the boiler, especially when the gauge glass is completely full or empty of water.



Figure 3 Water gauge fixed to boiler

If necessary, replace a gauge glass, even if the boiler must be shut down. That inconvenience is nothing compared with the damage that can result from a boiler being operated without a

functioning gauge glass. The connection lines to a gauge glass can become clogged and show normal water levels when water is low; thus, the piping connecting a gauge glass to a boiler should be cleaned and inspected regularly to ensure it is clear.

A boiler's fuel system, particularly the burner, requires periodic cleaning and routine maintenance. Failure to maintain equipment in good working order can result in high fuel costs, the loss of heat transfer, or a boiler explosion.

Pressure and temperature gauge side by side, just on the vessel and in the front so that the worker can see that the inside temperature has become atmospheric and it is safer to open the manhole cover. Better hinges of bottom filter cover and big size internal diaphragm plate on filter cylinder to restrict the sudden outflow of hot water, Periodical pressure and corrosion testing of the vessel and its safety devices. Quality welding joints with radiography test and Training to workers for safe operation and maintenance. Use of soft water without salts/chlorides and heating/cooling through heat exchanger are necessary to reduce the frequency of failure.

#### V. OTHER SAFETY FEATURES

All normal safety precautions should be followed when operating boilers, burners, and fuel systems.

*Manufacturer's Instructions* - Equipment manufacturer's instructions should be followed.

*Training* - Employees must be trained in safety prior to operation of the equipment. The training in safety should be a continuous process for the purpose of educating employees to recognize and to keep safety in their minds throughout their careers. For that regularly training program should be established and maintained.

*Housekeeping* - Good housekeeping is essential for safety and good plant operation. Poor housekeeping results in increased safety hazards. A clean and orderly environment will foster safety.

*Clothing and Protective Equipment* - Proper clothing should be worn at all times. Avoid loose clothing and jewelry. Protective equipment must be worn when necessary (i.e., helmets, respirators, ear plugs, goggles, gloves, safety shoes, etc.). Never operate rotating equipment, mechanically automated devices, or electrically and pneumatically operated control components unless guides, shields, or covers are in place.

*Hot Surfaces* - Many hot surfaces exist in a boiler area and even non-heated surfaces can become uncomfortably warm, therefore, employees, especially new employees, must be made aware of these conditions. Refractory and insulation are typically provided to encounter elevated surface temperatures in some installations. Care must be exercised to prevent burns and other thermal hazards when near the boiler. Never enter the boiler until an adequate cool-off period has been observed and the Owner's entry procedures have been completed.

*Lockout and Tagout Procedures* - Every plant should have a formalized lockout and tagout procedure that is strictly enforced. It is a method of protecting employees from accidental machine startup through proper locking and labeling of machines that are undergoing maintenance.



Figure 4 Lockout and Tagout

*Remote Starting of Equipment* - Much of the equipment in plants are started remotely and/or automatically without warning; therefore, employees must be alert to avoiding that equipment which can be started remotely. If work is to be done on any equipment, lockout and tagout procedures must be followed. Attach signs to equipment such as "DO NOT START - MEN AT WORK". Attach a similar sign on the equipment control panel.



Figure 4 Safety Sign



Figure 6 Safety Signs

*Fire and Explosion Hazards* - A fired boiler utilizes fuels which are flammable and potentially explosive. Extreme care should be exercised when making fuel-piping connections. Use the correct gasket, bolts, thread lubricants, and tightening torque to prevent leaks. It is recommended that drain valve and/or vent piping be channeled to safe locations. Valve packing should be periodically tightened and a rigorous leak check program be implemented as part of the Owner's preventative maintenance program.



Figure 5 Fire Safety Equipments

*Electrical Hazards* - Potentially hazardous voltages exist in control cabinets and electrically actuated control components. These components should only be serviced when system power is removed and only by qualified electrical or instrumentation servicemen.

*Unconventional Fuels* - Sometimes unconventional fuels need to be burned in boilers. When this is done, particular attention should be paid to the hazards that can result. These may from characteristics in the fuels, toxic chemicals in the fuel, and toxic chemicals produced through combustion. Persons knowledgeable in the use of such unconventional fuels should be consulted concerning the problems that may be encountered. Because of the wide variety and limited use, such fuels are not addressed in this manual.



VI. OBJECTIVES OF FACTORIES AND BOILERS INSPECTION DEPARTMENT

The industrial safety rules in relation to their causing damage to environment have been taken more seriously in our country. This has helped industries to become more environment friendly. The Central Boilers Board deals with the problems of both the users and manufacturers and takes policy decisions for the proper growth of the boiler manufacturing units in the country. The Board formulates the Indian Boiler Regulations incorporating the latest developments taking place in the Boiler Industry all over the world.

The following objectives of the factories and boiler inspection department are as below:

- To ensure safety of factory workers.
- To ensure health of factory workers.
- To ensure welfare of factory workers.
- To minimize factory accidents.
- To minimize boiler accidents.
- To eliminate child labour from factories.
- To monitor MAH (*MAJOR ACCIDENT HAZARD*) factories.
- To undertake survey under various laws.
- To enforce various laws, thereby ensuring welfare of the workers.

VII. CONCLUSION

In this study, safety precautions are defined to prevent the accidents in boiler industries. Hence, the designed safety education programs modules should be utilised by the industries for success in their accident control and prevention strategies, specifically to correct or alter the unsafe acts or mistakes committed by the workers at work, which are likely to result in accidents.

We can conclude that while there are safety devices to prevent dangerous boiler operating conditions from turning into disasters, only proper maintenance can prevent dangerous boiler operating conditions from occurring. The only way one can be confident control or safety devices are functioning properly is to perform required maintenance, testing, and inspection regularly.

VIII. TABLE.

OPERATING HAZARD	CAUSES	EFFECT	PREVENTION
Steam Leaks	Damaged or corroded pipes and/or other pressure parts	Severe burns.	Keep all joints and pipes tight.  Warn personnel of hazards of invisibility of superheated steam leaks.
Defective safety valves.	Obstruction between boiler and valve.  Valve damaged or corroded (internal). Lever tied down. Obstruction on valve outlet.	Will not lift to release excess pressure. Impose excess pressure on the boiler. Rupture the boiler. Cause loss of life and/or injury to personnel. Cause property damage.	Replace or repair safety valve.  Remove obstructions. Periodically test valve per ASME code.
Defective steam pressure gauges.	Broken gauge. Gauge is not in calibration. Blockage in line from boiler to gauge.  Gauge cock is closed. Multiple gauges not in	Gauge is not showing the correct pressure. Boiler may be under excessive pressure. Prevents operator from being aware of true operating	Calibrate gauge regularly.  Replace defective gauges.  Inspect gauge connection and piping to boiler for blockage and/or closed cock.

	agreement .	condition s.	
Low water level.	Defective low water cutoff. Low water cutoff bypassed. Improper water column blow down procedure . Equalizing lines restricted or plugged. Tampering with low water control. Defective boiler water feed system. Operator error. Defective or inoperative gauge glass.	Overheated boiler surfaces. Ruptured boiler. Loss of life and/or injury to personnel . Property damage.	Verify operation of boiler water feed system periodically. Prove low water cutoff operation periodically. Use proper water column blow down procedures. Train boiler operators. Replace defective low water controls. Inspect equalizing line

		Fire.	
Excessive negative pressure.	Flame out. Induced draft fan runaway.	Equipment damage resulting in personnel injury.	Maintain proper operation of control equipment. Do not bypass control equipment. Use procedures of NFPA 85G.
Activities related to cleaning.	Failure to observe safety procedures applicable to maintenance cleaning.	Potential injury or death to personnel.	Observe operating and maintenance instructions for maintenance cleaning. Observe all safety regulations and normal safety precautions. Provide personnel with protective clothing and equipment. Establish a routine procedure to clean. Report all unsafe conditions and/or unsafe practices.
Obstructed areas.	Poor housekeeping.	Potential injury to personnel.	Provide a safe means of access to all equipment and working places.

HAZARDS & PREVENTION IN BOILER

OPERATING HAZARD	CAUSES	EFFECT	PREVENTION
Leaking fuel safety shutoff valves.	Defective valve. Foreign matter under valve seat.	Fuel flows to the boiler. Uncontrolled ignition of fuel. Fireside explosion. Loss of life and/or injury to personnel. Boiler damage. Property damage.	Replace defective valves. Leak test and verify proper operation of valves periodically.

XI. REFERENCES

[1] [www.isplonline.com](http://www.isplonline.com)  
 [2] [http://dipp.nic.in/boiler\\_rules\\_updated/about.htm](http://dipp.nic.in/boiler_rules_updated/about.htm)  
 [3] <http://www.chemicalsafety.co.in/safety-law.pdf>  
 [4] <http://www.banksengineering.com/blrsafety.htm>  
 [5] M. Mahajan., “Industrial Engineering and Production Management”.  
 [6] Amitkumar Gupta, “Industrial Safety & Environment”.

# “Tool Life Prediction in Face Milling Machining Using Artificial Neural Networks”

A. Jignesh Parmar, B. Amit Patel, C. Chirag Patel

**Abstract**—Tool life is an important indicator of the milling operation in manufacturing process. Studies and analyses of milling process are usually based on three main parameters composed of cutting speed, feed rate and depth of cut. The aim of this study is to discover the role of these parameters in tool life prediction in milling operations by using artificial neural networks

**Index Terms**—Artificial neural networks, Face milling, Tool life prediction.

## I. INTRODUCTION

A milling machine is a power driven machine that cuts by means of a multi tooth rotating cutter. The mill is constructed in such a manner that the fixed workpiece is fed into the rotating cutter. Varieties of cutters and holding devices allow a wide range of cutting possibilities. Milling can be defined as machining process in which metal is removed by a rotating multiple-tooth cutter with each tooth removes small amount of metal in each revolution of the spindle. Mills are classified on the basis of the position of their spindle. The spindle operates in either a vertical or horizontal position. In modern industry the goal is to manufacture low cost, high quality products in short time. Predictive models of machining processes and tool life can be applied to help businesses gain a competitive edge. In this time of expanding global markets, it has become essential for manufacturers to improve process efficiencies, maintain stricter part tolerances, and enhance part quality. Furthermore, the motivation for using analytical tools for process optimization, rather than costly trial and error, has perhaps never been greater. Dynamic models of milling processes provide the ability to predict stable cutting condition and increases tool life for a large combination of process. The application of the Artificial Neural Networks (ANN) model for the modeling purpose in various different areas including machining is used very widely by researchers. A good review on cutting force control and tool wear monitoring in end milling can be found in [1]–[7]. Soichi and Takuya studied on a long-term control scheme of cutting forces to regulate tool life in end milling processes [8]. Indices based on milling force for tool wear in milling have

been investigated by Yan et.al [9]. Bhattacharyya et.al [10] investigated on cutting force-based real-time estimation of tool wear in face milling using a combination of signal processing techniques while tool wear in high speed milling using detection process approach has been done by Kious et.al [11]. Estimation of tool wear during CNC milling using neural network-based sensor fusion was implemented by Ghosh et.al [12]. On-line metal cutting tool condition monitoring using multi-layer perceptron neural networks was studied by Lister and Dimla [13]. Tool condition monitoring using artificial intelligence methods has been carried out by Balazinski et.al [14]. Cho and Ko estimated tool wear length in finish milling using a fuzzy inference algorithm [15]. Intelligent process supervision for predicting tool wear in machining processes was done Alique by and Haber [16]. Ning and Veldhuis[17] analyzed mechanistic modeling of ball end milling including tool wear. Prediction of flank wear by using back propagation neural network modeling when cutting hardened H-13 steel with chamfered and honed CBN tools is used by Ozel and Nadgir [18] Also tool cutting force modeling in ball-end milling using multilevel perceptron was implemented by Zuperl and Cus [19]. Rivero et.al [20] worked on tool wear detection in dry high-speed milling based upon the analysis of machine internal signals and Orhan et.al [21] evaluated tool wear by vibration analysis during end milling of AISI D3 cold work tool steel with 35 HRC hardness. Furthermore Tool wear perdition from acoustic emission and surface characteristic via an artificial neural network has been carried out by Wilkinson and Reuben [22]. By considering the abilities and limitations of above approaches for the tool life and tool wear monitoring, the focus of this study is finding the relation of cutting parameters (feed rate, depth of cut and spindle speed) on tool life and illustrate the capability of (ANN) to predict and modeling tool life.

## II. EXPERIMENTAL SETUP

### A. Milling machine

Experiments were carried out at the university of Applied Science and Technology on a milling machine as shown in Fig 1. The experiments were

conducted in a ZXX6350ZA vertical axis milling machine using tool life 20mm diameter cutter mill with 4 cutter inserts (Fig 2). The cutter mill was made by high speed steel-E (HSS-E). Work-pieces used for this experiment was Al 7075.



Fig 1. Milling machine



Fig 2. End mill

*B. The Taguchi design of experiments method*

The most efficient method of experimental planning is Design of Experiments (DOE) using the Taguchi approach, which was adopted in this paper. (DOE) incorporates the orthogonal arrays, developed by Taguchi, to successfully design and conduct fractional factorial experiments that can collect all the statistically significant data with the minimum possible number of repetitions. Full factorial experiments are conducted or one factor at a time strategies are followed. The former cannot be implemented when there are too many factors under consideration because the number of repetitions required would be prohibitive, from a time and cost viewpoint. In this experiment each parameter has 5 levels which are so degree of freedom (DOF) for each parameter is 4. The total calculated (DOF) is  $4 \times 4 \times 4 \times 4 = 1024$ . After determining the number of (DOF), the next step is to choose suitable orthogonal

array. In Taguchi (DOE), orthogonal array must be more than or equal to the (DOF) of design parameters. So the best orthogonal array by using Taguchi DOE is L25 ( $5 \times 5$ ). This array contains 25 repetitions (Table 1).

TABLE I TAGUCHI (DOE) IN THIS WORK

No	Spindle speed (rpm)	Feed rate (mm/min)	Depth of cut (mm)	Tool life measured during process (min)
1	95	22	0.2	264
2	95	98	0.4	27
3	95	132	0.6	23
4	95	200	0.8	18
5	95	360	1	12
6	360	22	0.4	554
7	360	98	0.6	455
8	360	132	0.8	398
9	360	200	1	316
10	360	360	0.2	303
11	565	22	0.6	612
12	565	98	0.8	462
13	565	132	1	402
14	565	200	0.2	387
15	565	360	0.4	276
16	950	22	0.8	180
17	950	98	1	157
18	950	132	0.2	350
19	950	200	0.4	367
20	950	360	0.6	153
21	1500	22	1	160
22	1500	98	0.2	163
23	1500	132	0.4	140
24	1500	200	0.6	113
25	1500	360	0.8	46

**III. MODELING OF THE TOOL LIFE BY ARTIFICIAL NEURAL NETWORKS**

In this section, based on the experimental data, the tool life modeling and prediction is carried out by using multi layer perceptron (MLP) neural networks. MLP neural network is one of the most popular supervised (ANN) which has the ability to solve nonlinear problems. 25 examples have been used for the off-line training and performance checking of the proposed model. As shown in Fig 3, a four-layer (MLP) is used, including 3 inputs (i.e. cutting speed, feed rate and depth of cut), 2 hidden layers containing 5 neurons, and an output layer with a single neuron (i.e. tool life),  $3 \times 3 \times 2 \times 1$ .

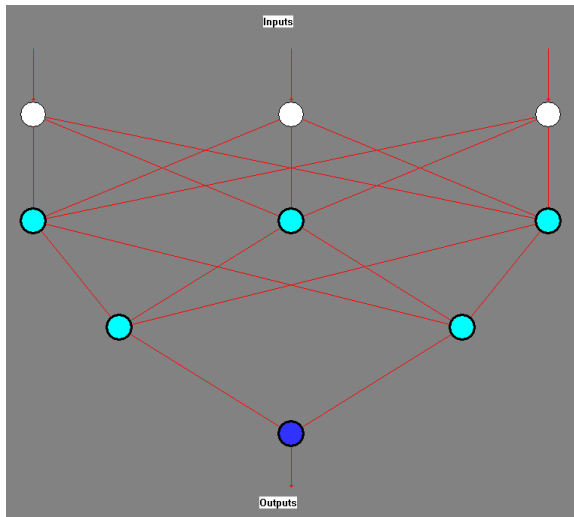


Fig 3. Proposed perceptron neural networks

Hidden layers must have equitable covering rate on learning data therefore, the best architecture and parameters of the (MLP) model are chosen through several tests which are not presented in this paper. Sigmoid, Gaussian and Hyperbolic Scant transfer functions have been applied for neurons of hidden and output layers, respectively. The obtained Root Mean Squared (RMS) error and correlation between the estimated and experimental datamwere 0.005575 and 0.96966, respectively (cf. Figures 4, 5 and 6).

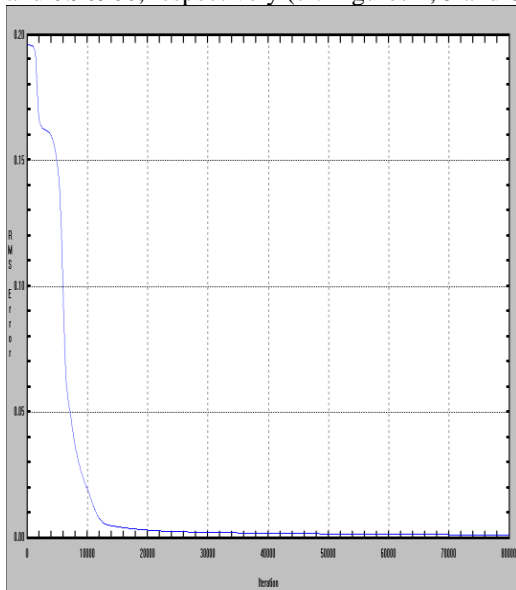


Fig 4. RMS error

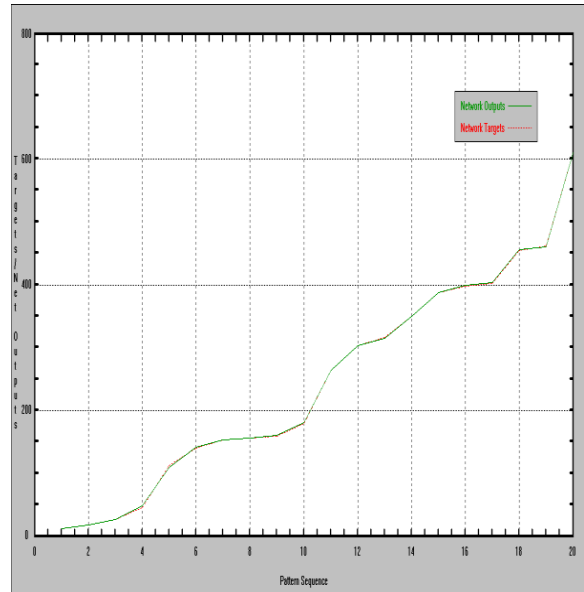


Fig 5. Correlation and curve fitting diagram for train data

Figure 6 illustrates correction coefficient versus iteration and after 10000 iteration correction coefficient reach to peak. As shown in Figures 4 to 6, the proposed (MLP) neural network has provided proper modeling results. Indeed, this method can be reliably and successfully used for modeling of the tool life modeling.

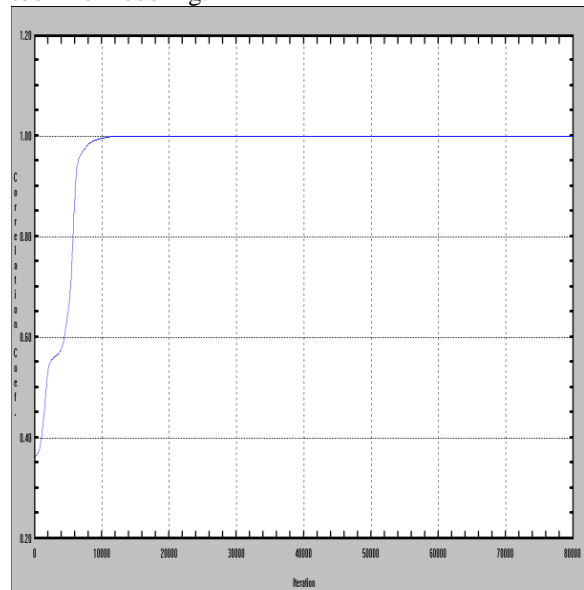


Fig 6. Correction coefficient versus iteration

**4. RESULTS**

To verify the accuracy of the obtained results, several experimental tests have been implemented. Table 2, illustrates good performance of (ANN) to predicting tool life in milling process. The obtained correlation between the estimated and experimental data was 0.94966. (cf. Figures 7 and 8).

TABLE II RESULTS OF (ANN) PREDICTION AND EXPERIMENTAL MEASUREMENT

No	Spindle speed (rpm)	Feed rate (mm/min)	Depth of cut (mm)	Tool life measured during process (min)	Tool life predicted by ANN (min)	Error %
1	95	132	0.6	23	26.43797	-0.14948
2	360	22	0.4	554	582.8904	-0.05215
3	565	360	0.4	276	276.3392	-0.00123
4	950	200	0.4	367	371.4312	-0.01207
5	1500	98	0.2	163	165.5439	-0.01561

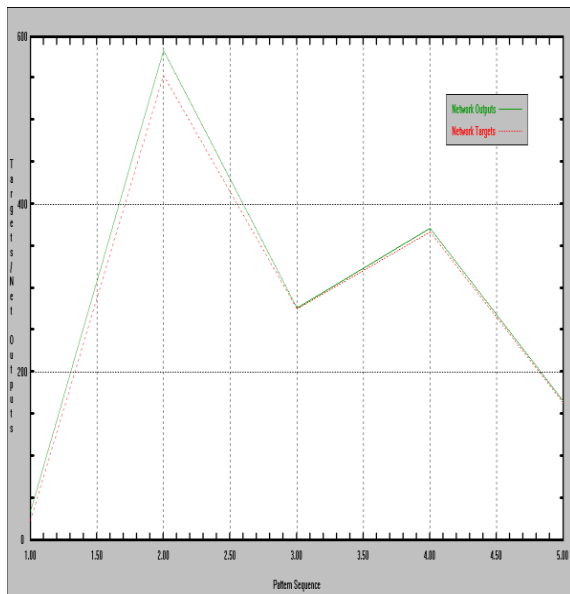


Fig 7. Correlation and curve fitting diagram for test data

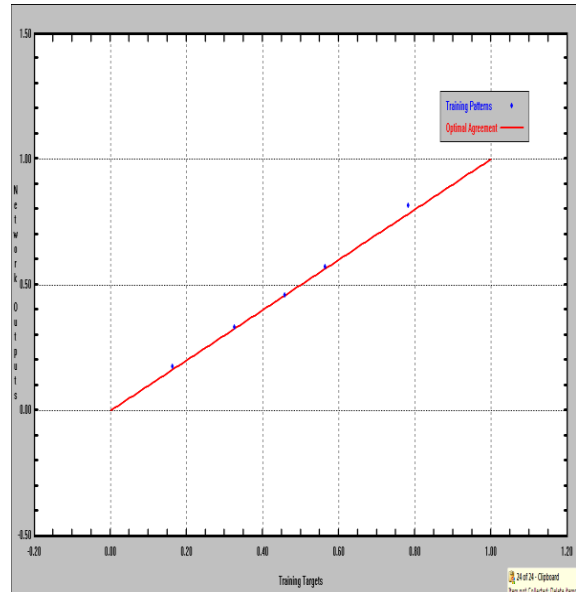


Fig 8. Correlation and scattering diagram for test data

**V. CONCLUSION**

In this study, (ANN) for modeling and predicting tool life in milling parts made of Aluminum (7075) material was developed. Given the accuracy that was achieved it is safe to conclude that all the significant factors were included in the (DOE) process. The research in the present paper can be extended towards three different steps. The first step is using Taguchi (DOE) and different combinations of cutting parameters for building database. The second step is modeling tool life by using (ANN). Third step is validation by carrying out the experimental tests. In generating the (ANN) model statistical (RMS) was utilized. The accuracy error was found to be insignificant (3.034%). It was found that (ANN) prediction correlates very well with the experimental results. Finally the correlation for training and test was obtained 0.96966 and 0.94966 respectively and mean square error was calculated 3.1908% for test data.

**REFERENCES**

- [1] SY. Liang, RL. Hecker, RG. Landers “Machining process monitoring and control: the state-of-the-art,” *Trans ASME Journal of Manufacturing Science Engineering*, 126–2, 2004, pp. 297–310.
- [2] AG. Ulsoy, Y. Koren “Control of machining processes,” *ASME Journal Dynamic System Measurement Control*, 115, 1993, 301–8.
- [3] Altintas Y. “Manufacturing automation: metal cutting mechanics,” machine tool vibrations, and CNC design, Cambridge University Press; 2000, pp 100-103.
- [4] A. Matsubara, S. Ibaraki “Monitoring and control of cutting forces in machining processes: a review,” *International Journal of Automation Technology* 2009, pp 445–456.
- [5] A. Matsubara, “Current status and trends of monitoring and control technology in machining process,” *Journal Society Instrument Control Engineering*, Vol 41 2002, pp 781–786.
- [6] E. Dimla, S. Dimla, “Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods,” *International Journal of Machine Tools & Manufacture*, Vol 40, 2000, pp 1073–1098.
- [7] P.W. Prickett, C. Johns, “An overview of approaches to end milling tool monitoring,” *International Journal of Machine Tools & Manufacture*, Vol 39, 1999, pp 105–122.
- [8] S. Ibaraki, T. Shimizu, “A long-term control scheme of cutting forces to regulate tool life in end milling processes,” *Precision Engineering*, Vol 34 2010, pp 675–682.
- [9] W. Yan, Y.S. Wong, K.S. Lee, T. Ning, “An investigation of indices based on milling force for tool wear in milling,” *Journal of Materials Processing Technology*, Vol 89, 1999, pp, 245–253.
- combination of signal processing techniques,” *Mechanical Systems and Signal Processing*, Vol 21, 2007, pp 2665–2683.
- [11] M. Kiouss, A. Ouahabi, M. Boudraa, R. Serra, A. Cheknane, “Detection process approach of tool wear in high speed milling,” *Measurement*, Vol 43, 2010, pp 1439–1446.
- [12] N. Ghosh, Y.B. Ravi, A. Patra, S. Mukhopadhyay, S. Paul, A.R. Mohanty, A.B. Chattopadhyaya, “Estimation of tool wear during CNC milling using neural network-based sensor fusion,” *Mechanical Systems and Signal Processing*, Vol 21, 2007, pp 466–479.
- [13] D.E. Dimla Sr. a, P.M. Lister, “On-line metal cutting tool condition monitoring.II: tool-state classification using multi-layer perceptron neural networks,” *International Journal of Machine Tools & Manufacture*, Vol 40, 2000, pp 769–781.
- [14] M. Balazinski, E. Czogala, K. Jemielniak, J. Leski, “Tool condition monitoring using artificial intelligence methods,” *Engineering Applications of Artificial Intelligence*, Vol 15 , 2002, pp 73–80.
- [15] T. J. Ko, D. W. Cho, “Estimation of tool wear length in finish milling using a fuzzy inference algorithm,” *Wear*, Vol 169, 1993, pp 97- 106.
- [16] R. E. Haber, A. Alique, “Intelligent process supervision for predicting tool wear in machining processes,” *Mechatronics*, Vol 13, 2003, pp 825–849.
- [17] L. Ning, St. C. Veldhuis, “Mechanistic Modeling of Ball End Milling Including Tool Wear,” *Journal of Manufacturing Processes*, Vol. 8, 2006, pp 21-28.
- [18] T. Ozel, A. Nadgir, “Prediction of flank wear by using back propagation neural network modeling when cutting hardened H-13 steel with chamfered and honed CBN tools,” *International Journal of Machine Tools & Manufacture*, Vol 42, 2002, pp 287–297.
- [19] U. Zuperl, F. Cus, “Tool cutting force modeling in ball-end milling using multilevel perceptron,” *Journal of Materials Processing Technology*, Vol 153–154, 2004, pp 268–275.
- [20] A. Rivero, L.N. López de Lacalle, M. L. Penalv, “Tool wear detection in dry high-speed milling based upon the analysis of machine internal signals,” *Mechatronics*, Vol 18, 2008, pp 627–633.
- [21] S. Orhan, A. Osman Er, N. Camus-cu, E. Aslan, “Tool wear evaluation by vibration analysis during end milling of AISI D3 cold work tool steel with 35 HRC hardness,” *NDT&E International*, Vol 40, 2007, pp 121–126.
- [22] P. Wilkinson, R. L. Reuben, “Tool wear perdition from acoustic emission and surface characteristic via an artificial neural network,” *Mechanical Systems and Signal Processing*, 13(6), 1999, pp 955-966. [10] P. Bhattacharyya, D. Sengupta, S. Mukhopadhyay, “Cutting forcebased real-time estimation of tool wear in face milling using a

# Computer Simulation of Triple Pressure Combined Cycle Power Plant

A. Kautilya D. Thummar, B. Alkesh M. Mavani

**Abstract**—Computer simulation is carried out of the combined cycle power plant (having triple pressure bottoming cycle) by computer software GateCycle™ to predict the performance of the combined cycle power plant. The performance of various heat exchangers like economizers, super-heaters, evaporators etc. has been determined by using the concept of approach temperature and pinch delta temperature. The results obtained from the simulation of the system are close to the actual performance of the combined cycle power plant.

**Keywords:** Combined Cycle Plants, GateCycle™, Performance, Simulation, Triple pressure Cycle.

## Nomenclature:

CCPP	Combined Cycle Power Plant
HP	High Pressure
IP	Intermediate Pressure
LP	Low Pressure
LHV	Lower Heating Value
HX	Heat Exchanger

## INTRODUCTION

With the twin crises of energy resources depletion and pollution increasingly engulfing our civilization, it has become very crucial to develop more efficient and less polluting power plants, which are capable of effectively utilizing fuels like coal and natural gas, etc. Due to the reason combined cycle power plants has recently received worldwide interest because of their high energy conversion efficiency and low emission capability. Combined cycles have the higher thermal efficiency as well as output power in comparison with gas turbine cycle and steam cycles. Higher efficiencies of combined cycle power plants (CCPPs) compared to Brayton or Rankine cycles have made them quite attractive for power generation. The exhaust from the gas turbine is used to generate steam

which drives a conventional bottoming steam plant working on Rankine cycle.

The design of combined cycle power plant is inherently complex because of involvement of two different power cycles. Therefore, a proper distribution of the power between two cycles is necessary between them. Therefore a need for developing computer simulation techniques which would enable evaluation of various possible design options and also permit prediction of off-design performance of the system.

This paper presents the simulation procedure which has been developed for predicting the performance of a typical CC power plant involving a gas turbine coupled to a triple-pressure bottoming cycle through a heat recovery steam generator. The procedure has been validated by comparing its prediction to the rated performance of a typical 384 MW, module combined cycle power plant, which consists of one gas turbine and three steam turbines.

## COMPUTER SIMULATION STRATEGY

By computer simulation of the system, it predicts the operating conditions of the system (temperatures, pressures, energy, mass flow rates etc.) at various components with the help of mass and energy balances. Therefore, the availability of performance characteristics of the various components constituting the system is a pre-requisite for the system simulation. The strategy of system simulation is strongly dependent on the manner in which the characteristics of various components are available. For the purpose of simulation, the output variables are the functions of the known input variables. Thus it is possible to choose the input and output variables judiciously to arrive at an optimal simulation strategy.

## CASE DESCRIPTION

As we know, gas-steam cycle i.e. combined cycle plants uses the waste heat from the gas turbine to produce steam in bottoming steam. The steam is produced in the bottoming cycle by utilization of the HRSG's. Higher the pressure level of HRSG's the better the utilization of the heat which was being previous wasted. The triple pressure system which is been discussed here is better for this purpose. The

---

A.Mechanical Engineering Department,  
L.D.R.P. Institute of Technology & Research,  
Gandhinagar-382015. India  
Email: [thummar.kautilya@gmail.com](mailto:thummar.kautilya@gmail.com)

B.Associate Professor, Mechanical Engineering Department,  
L.D.R.P. Institute of Technology & Research,  
Gandhinagar-382015, India.  
Email: [ammtimes@gmail.com](mailto:ammtimes@gmail.com)



typical arrangement of the system being simulated is shown in the figure 1.

The arrangement for triple pressure is by first placing the high pressure superheater, evaporator, economizer and then by the heat exchanger which works as reheater. The intermediate superheater, evaporator, and economizer are placed then after. Last stage of the triple pressure is the low pressure evaporator and superheater.

The exhaust gases leaving gas turbine having high temperature first's superheats the steam in the HP SPH2 before being expanded in the HP ST 1. The exhaust from HP turbine is mixed with steam coming from the IP SPH 2 and is reheated in the heat exchanger RHT. This mixture is expanded in the IP ST. the exhaust from the IP ST is again expanded in the LP ST. The steam from LP turbine is extracted at low pressure and send to the deaerator through condenser. Before going to the deaerator the saturated steam is being heated in the heat exchanger HX 1 to raise the temperature of the saturated steam leaving the condenser. The steams with low-pressure and low-temperature will condensate in the condenser to saturated water. The saturated water out of the condenser is mixed with the steam inside the deaerator, resulting water. The saturated water leaving the deaerator is pumped by using three BFP for high, intermediate and low pressure BFP. The saturated water out of deaerator is pumped by low pressure BFP to the evaporator through LP drum, where the phase change takes place by latent heat of vaporization and saturated water is converted in to the steam. This steam is then again sent back to the LP drum for the recirculation, this is necessary as enough energy is need to available for deaerator to fully deaerate the water. Now steams from LP drum goes to low pressure superheater 1 and 2, and expanded into the LP ST by getting mixed with the exhaust from the IP turbine.

The saturated water from the deaerator is pumped to intermediate pressure by IP BFP and goes to the IP economizer where it is heated up. The heated water then goes to IP drum and then to IP evaporator where it is partially superheated. The partially superheated steam flows back to IP drum, from there to the superheater to get superheated. The superheating takes place in two stages by two super heaters. The superheated steam is then expanded in to the IP turbine.

Now the saturated water from the deaerator is pumped to high pressure by HP BFP and goes to HP economizer where it gains heat. The heated water is partially superheated in the HP evaporator by change in phase by gaining the latent heat of vaporization. The partially superheated steam is superheated inside the HP superheater and expanded into the HP turbine.

Figure 14 shows the total heat release diagram by the exhaust gases to the various components in the combined

EQUIPMENTS	VALUES	UNIT
<b>Ambient conditions</b>		
Temperature	302.2	°K
Pressure	1.01	bar
<b>GT</b>		
Turbine Exhaust Gas Flow	622.25	Kg/s
Turbine Exhaust Gas Temperature	885	°K
Gas Inlet Temperature	458	°K
Gas Inlet Pressure	35	bar
<b>ST</b>		
HP ST Exit Pressure	35.6	bar
HP ST Isentropic Efficiency	87%	
IP ST Exit Pressure	8.3	bar
IP ST Isentropic Efficiency	87%	
IP ST Exit Pressure	0.08	bar
IP ST Isentropic Efficiency	87%	
<b>HRSR</b>		
HP Economizer 2 Approach	283	°K
HP Economizer 1 Approach	283	°K
HP Evaporator Pinch	283	°K
HP Superheater 1 Approach	283	°K
HP Superheater 2 Approach	283	°K
IP Economizer Approach	283	°K
IP Evaporator Pinch	283	°K
IP Superheater 1 Approach	284.11	°K
IP Superheater 2 Approach	284.11	°K
LP Evaporator Pinch	283	°K
LP Superheater 1 Approach	283	°K
LP Superheater 2 Approach	283	°K
<b>DEAERATOR</b>		
Deaerator Operating Pressure	3.44	bar
<b>CONDENSOR</b>		
Condenser Operating Pressure	0.077	bar

cycle system.

INPUTS GIVEN

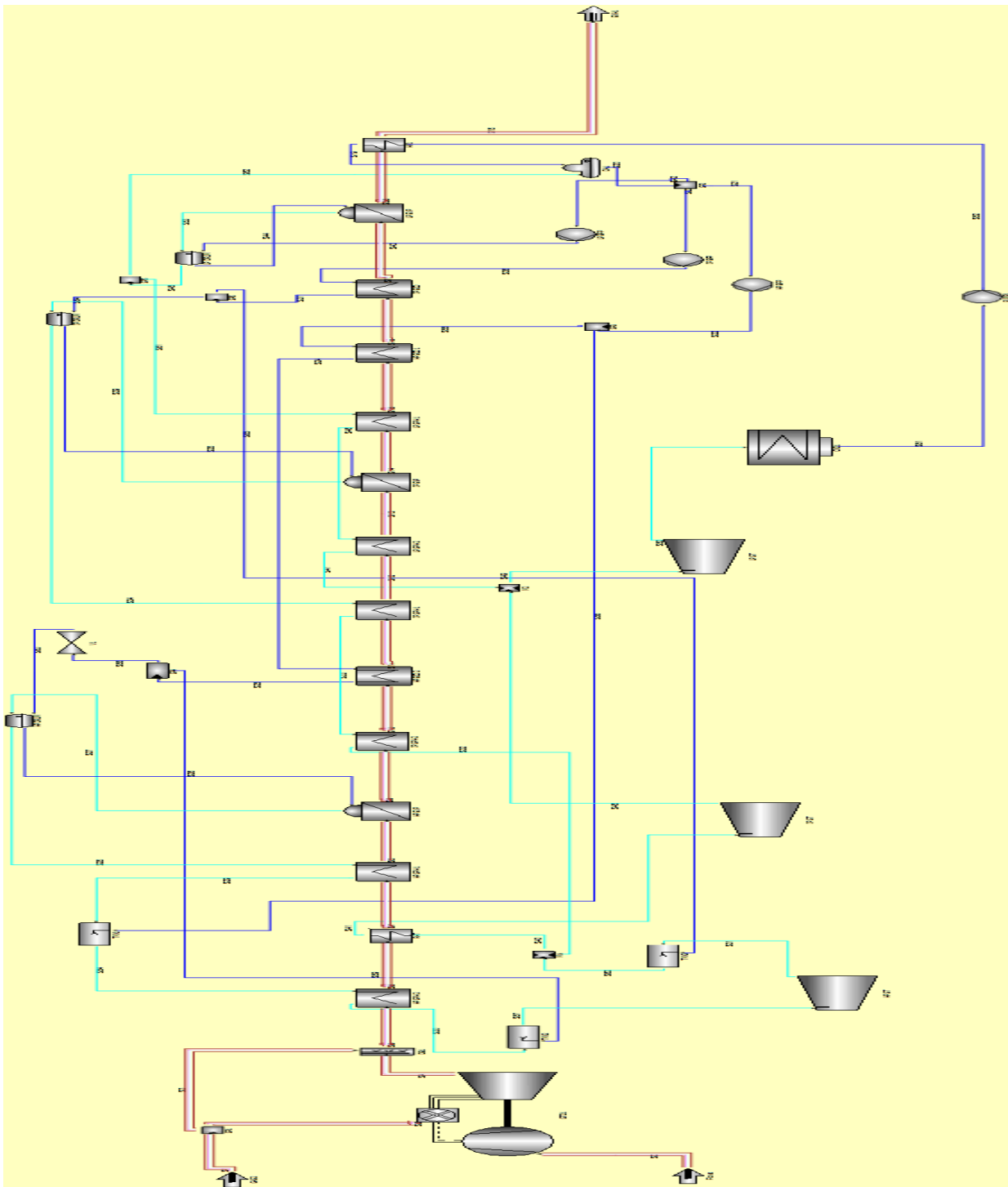


Fig. 1

GateCycle™ model of triple pressure CCPP

**RESULTS OBTAINED**

The simulations carried out of the above described system with required inputs, the results which are obtained are as follows.

Table 2:

Variable	VALUES	Unit
Net Cycle Power	375.4487	MW
Net Cycle Lower Heating Value (LHV) Efficiency	54.3032	
Net Cycle Lower Heating Value (LHV) Heat Rate	1.8414	kJ/kW-sec
Net Gas Turbine Power	239.58	MW
GT Shaft Power	243.2284	MW
GT Generator Losses	3.6484	MW
GT Simple-Cycle Lower Heating Value (LHV) Efficiency	36.4046	
Net Steam Cycle Power	135.8687	MW
ST Shaft Power	140.5701	MW
ST Generator Losses	2.8114	MW
Steam Cycle BOP Losses	1.89	MW
ST Generator Output	137.7587	MW
Adjusted Cycle Lower Heating Value (LHV) Efficiency	54.3032	
Adj. Cycle Lower Heating Value (LHV) Heat Rate	1.8414	kJ/kW-sec

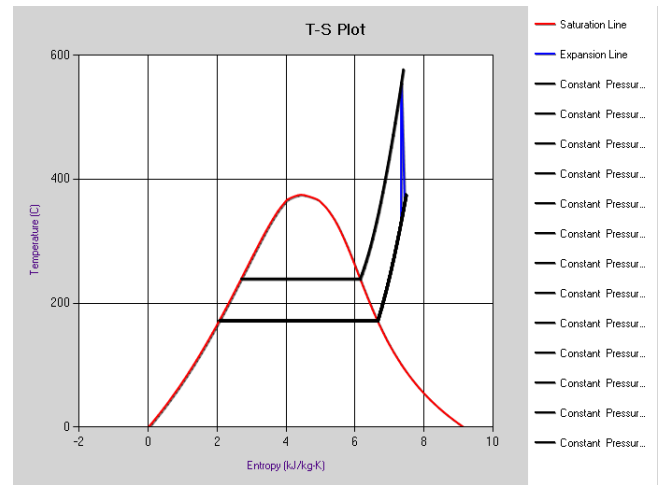


Fig. 3 T-S Diagram IP Turbine

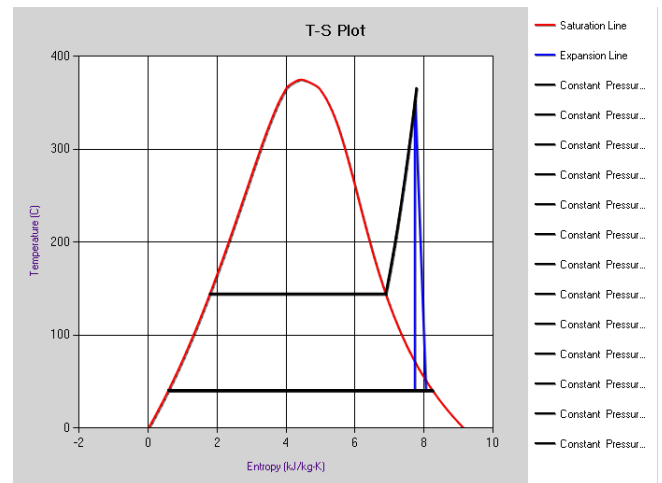


Fig. 4 T-S Diagram LP Turbine

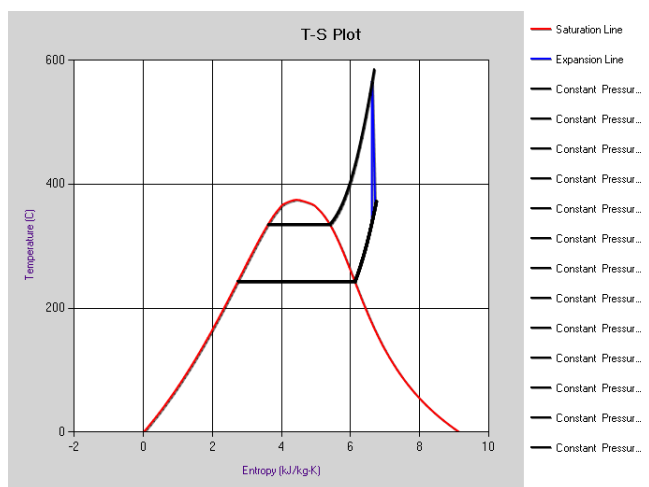


Fig. 2 T-S Diagram HP Turbine

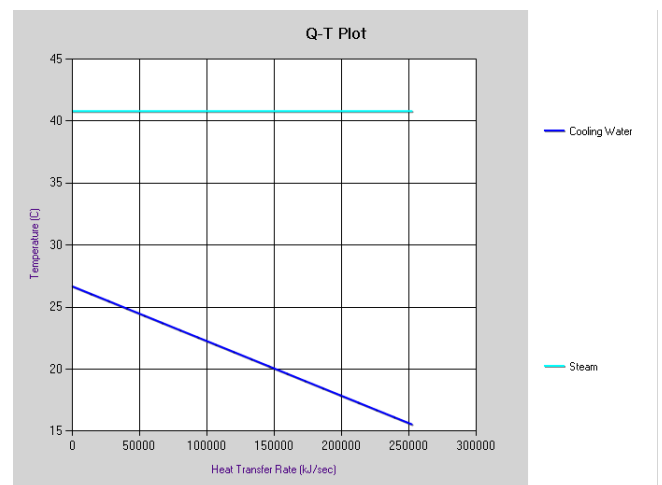


Fig. 5 Q-T Diagram Condensator

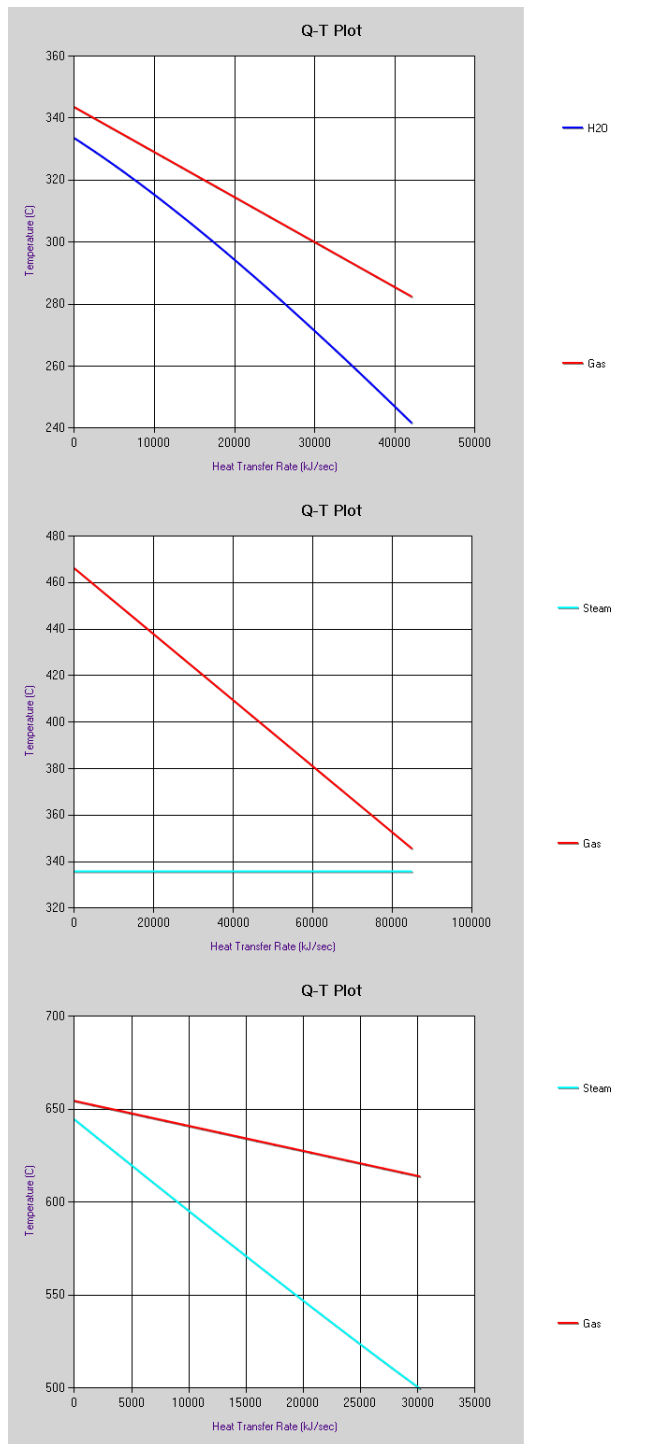


Fig. 6,7,8 Q-T Diagrams for HP Economizer, Evaporator & Superheater

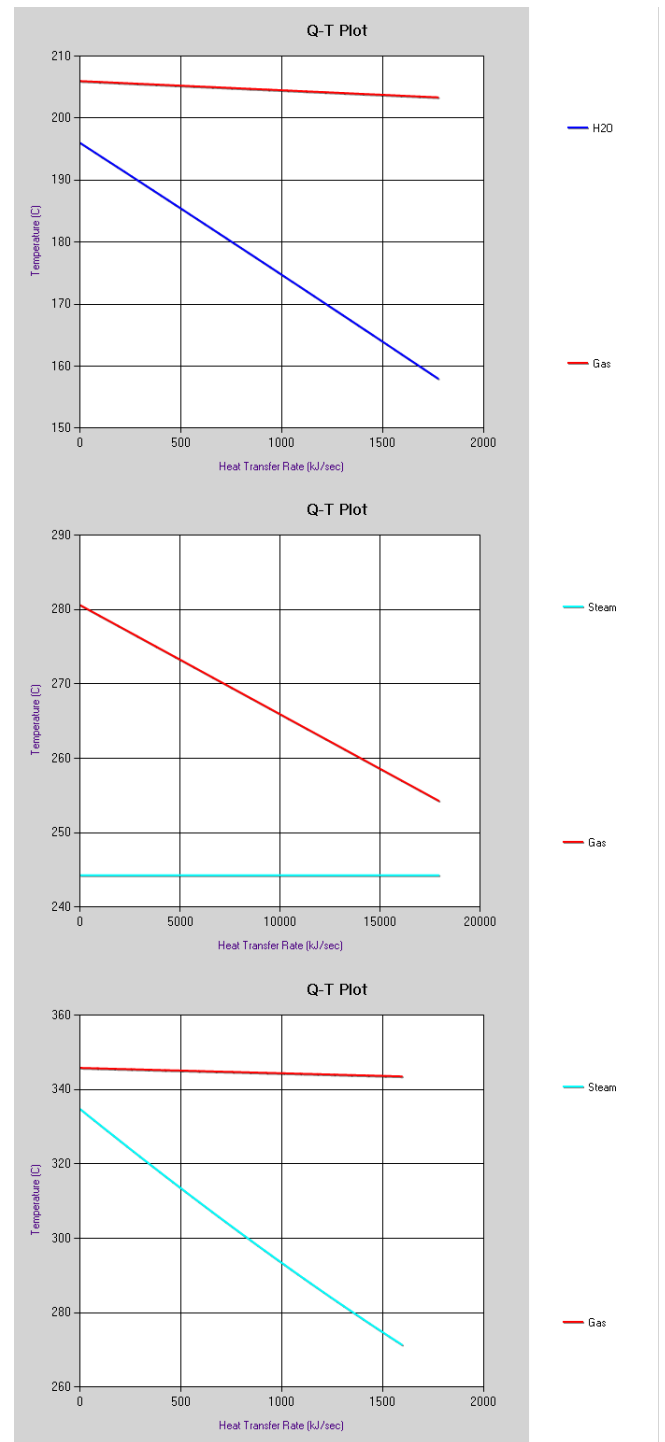


Fig. 9, 10, 11 Q-T Diagrams for IP Economizer, Evaporator & Superheater

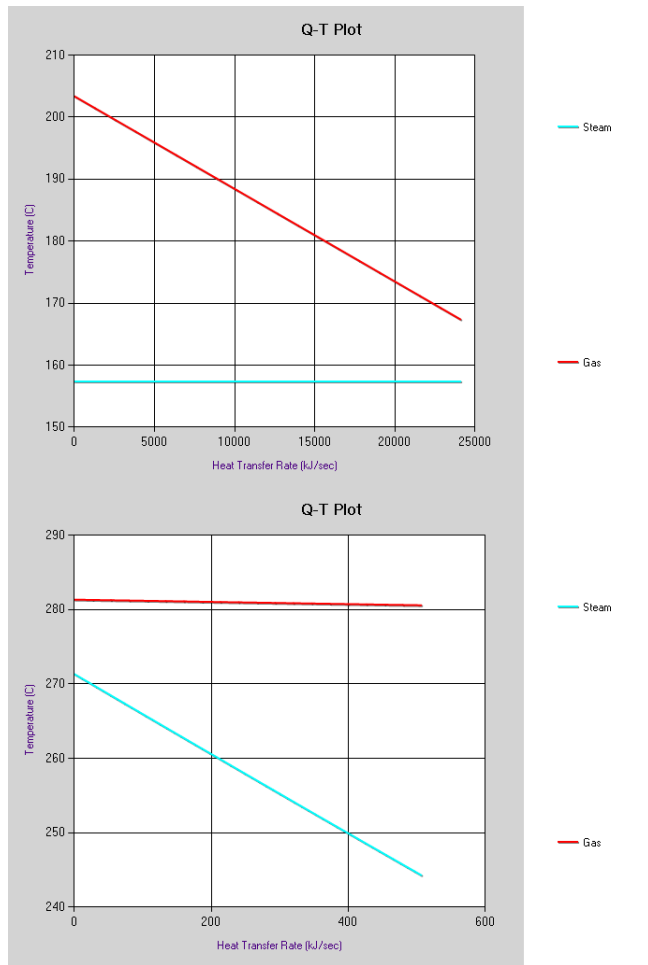
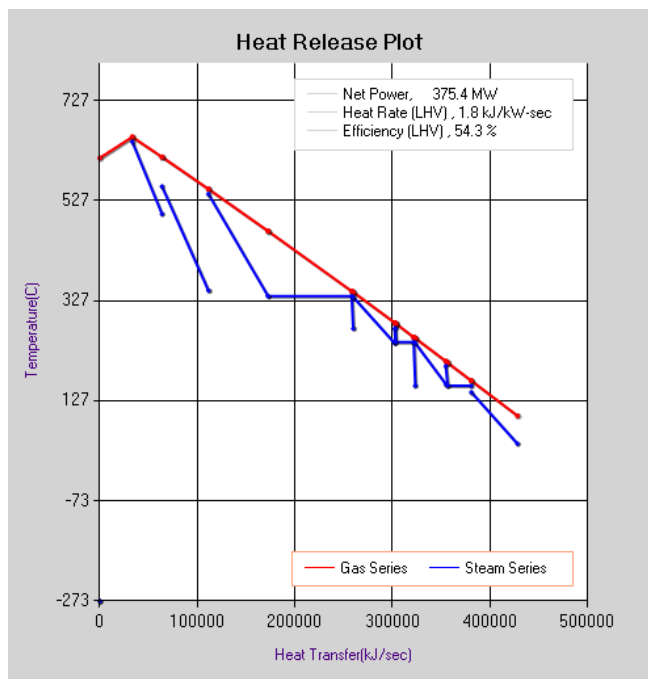


Fig. 12, 13 Q-T Diagrams for HP Evaporator & Superheater



CONCLUSIONS

The conclusion that can be derived by the simulation carried out is shown in the Table 2. The net efficiency of the plant come 375.4487MW which is very near to the predicated power expected to produce. The plant was originally designed to produce power of 384 MW. The variation is about 2% which is acceptable. The graph shows the expansion of steam in turbines. The power produced by LP turbine is highest as it can be seen in the Fig 4. The overall heat transfer of the exhaust gases to the various components in HRSG’s (Heat Recovery Steam Generators) is shown in the Fig 14.

The efficiency of the simulated system came 54.3%, which is very high when compared with the individual efficiency of the gas turbine and the steam turbine. Hence we can say that the overall efficiency of the combined cycle plant is higher.

ACKNOWLEDGEMENT

The author acknowledges L&T-S&L for providing the required data to carry out the simulation of the system. The authors are also thankful to L&T-S&L which allowed to use GateCycle™ software. The authors also greatly appreciate Professor A.M. Mavani and are very thankful for all his help, guidance, suggestions and advice throughout this study. They would also like to thank Mr. Salil Kelkar of L&T-S&L, without whom this simulation process cannot be carried.

REFERENCES

NicoWoudstra, “Thermodynamic evaluation of combined cycle plants”, *Energy Conversion and Management* Vol 51 (2010), pp. 1099–1110

P. K. Nag, “Thermodynamic analysis of a coal-based Combined cycle power plant”. *Heat Recovery Systems & CHP* Vol. 15, No. 2, pp. 115-129, 1995

B. Seyedan, “Computer Simulation of acombined cycle power plant”, *Heat Recovery Systems & CHP* Vol. 15, No. 7, pp. 619-630, 1995

GEEnter Software LLC, Menlo Park, CA94025, USA, GateCycle™ (Version 6.1.0.0), 2010.

C. Shah “Simulation and Computation of combined cycle power system through HRSG’s heat exchangers arrangement and parameters for exergy and cost.”

[www.ge.com/energy](http://www.ge.com/energy)

Fig. 14 Heat Release Plot of all Components in the system

# Design, Development & Fabrication of Small Economical Downdraft Venturi Type Gasifier.

A.M N Patel, B. Prof. K P Trivedi, c. Prof. J M Patel,

**Abstract**— Increasing industrialization and motorization of the world led to a steep rise in demand of petroleum products. Petroleum based fuels are stored fuels in the earth. There are limited reserves of these stored fuels and they are irreplaceable. These finite resources of petroleum are highly concentrated in certain regions of the world has given rise to uncertainty in its supply and price and is impacting growing economies like India, importing 80% of the total demand of crude. Much of a research work all round the world is for alternate fuels & most of them pertaining towards non-renewable energy sources. Biodiesel has shown a great potential in the liquid fuel region. In the gas region natural gas & liquefied petroleum gas has played a main role but increasing prices of the same has been an alarming signal. Another source is biomass which has potential to produce gas called producer gas or syngas from gasification. This attempt involves design, fabrication, developing & testing a small downdraft biomass gasifier at L D College of engineering with mild steel pipes, sheets & other fittings to test for wood chips, pallets & bagasse as feed material. The air quantity which is the primary requirement of the gasifier as a controlling parameter is studied basic down draft venturi type gasifier for maximum tar cracking and purity of gas generation with different conditions. Paper involves literature survey for different down draft gasifier and based on which drawings were prepared. Design involves venturi type central nozzle with different throat diameter of 44, 76 & 100mm with throat inlet cone angle 45°, 60° & 70° and different nozzle height of 80, 100 & 120 mm. Gas composition, moisture content and consumption of biomass feed stock, temperature inside the reaction zone, primary air flow & exit temperature of the producer gas can be measured. Venturi type design found to be the best optimized design for down draft gasifier & for dual fuel engine performance as it ensures optimized gas generation & tar cracking.

**Keywords**— Downdraft gasifier, Venturi, Throat, Inlet cone angle, Nozzle Height.

## I. INTRODUCTION

Scanning through the published literature it is found that most of the published design approaches are based upon empirical rules for which practically no explanation is available. There is no unanimity regarding these design rules, so much so that it has been called as an art rather than science. However, there do exist some sincere efforts towards documentation of design rules for gasifiers based on simplified theoretical calculations and evaluation of existing gasifier designs. There is Evidence of Imbert design having been used most extensively. The design recommendation regarding the no. of air nozzle and their orientation is mostly governed by the consideration of achieving a uniform temperature of oxidation zone [1].

A. Master of Engineering, ICE & Automobile Engineering, L D College of engineering, Ahmedabad,

B. Associate Prof., L D College of engineering, Ahmedabad,

C. Associate Prof., L D College of engineering, Ahmedabad. mehumnp@yahoo.com

Wood gasifiers played an important role in the past in the substitution of oil-based fuels in internal combustion engines, but fell into disuse after the Second World War because of their economic and technical disadvantages as compared with relatively inexpensive imported fuels. It appears that interest in gasification research correlates closely with the relative cost and availability of liquid and gaseous fossil fuels [3]. In today's scenario, the fossil fuel prices are so high & are increasing day by day. Situation is pointing towards development of alternate source of energy with lower environmental emissions. Gasification paper tried to focus on outlook of development of small economical venturi type gasifier refereeing to its design. Producer gas, the gas generated when wood, charcoal or coal is gasified with air, consists of some 40 per cent combustible gases, mainly carbon monoxide, hydrogen and some methane. The rest are non-combustible and consists mainly of nitrogen, carbon dioxide and water vapour. The gas also contains condensable tar, acids and dust. These impurities may lead to operational problems and abnormal engine wear so possibilities of using it with the type of engine can be checked as per the quality of gas generated. The main problem of gasifier system design is to generate a gas with a high proportion of combustible components and a minimum of impurities [3].

In the simple word it can be said that with low tar content, other impurities & higher calorific value, gas can be used to replace diesel fuel up to amountable percentage in diesel engine. As the gasification process is done in controlled air environment, the tar cracking & maximum pure gas generation are two criterias on which most designs are based. Possibilities of using producer gas with different types of engines shows Spark ignition engines, normally used with petrol-or kerosene, can be run on producer gas alone or converted to full producer gas operation by lowering the compression ratio and the installation of a spark ignition system. Another possibility is to run a normal unconverted diesel engine in a "dual fuel" mode, whereby the engine draws anything between 0 and 90 per cent of its power output from producer gas, the remaining diesel oil being necessary for ignition of the combustible gas/air mixture. The advantage of the latter system lies in its flexibility: in case of malfunctioning of the gasifier or lack of biomass fuel, an immediate change to full diesel operation is generally possible [3].

## II. DESIGN CRITERIA FOR DOWNDRAFT GASIFIER

While working with this part of gasifier, it was found that gasifier developed till now by the industry are mostly patented items & hence very less part or no part of design is published in the public domain for design of gasifier. P. Basu [5]

describes characteristics of fuel (wood) for fixed-bed down draft gasifier as shown in table I.

TABLE I  
CHARACTERISTICS OF FIXED-BED GASIFIERS- FUEL (WOOD)

Moisture wet basis (%)	25 max
Dry-ash basis (%)	6 max
Ash melting temperature (°C)	>1250
Size (mm)	20–100
Application range (MW)	1-2
Gas exit temperature (°C)	700
Tar (g/Nm <sup>3</sup> )	0.015–3.0
Gas LHV (MJ/Nm <sup>3</sup> )	4.5–5.0
Hot-gas efficiency (%)	85–90
Turn-down ratio (–)	3-4

As per the literature survey & visit to M/s. Ankur Scientific, Savali, Baroda & M/s. Suntrack energy Pvt. Ltd., Ahmedabad & guidance of different experts, following criteria found to be most critical in downdraft throat type gasifier as per the design point of view. The list includes, the size of throat diameter, size of air nozzle, required air quantity for gasification, point of air entry (i.e. height & circumferential diameter), hearth load, required size of hopper & particle size & shape for better fluidity, bulk density of feed material, quality of feed material, size of reduction zone etc.

A. Recommendation by different researchers

Two different designs studied under the review. The design recommendation regarding the no. of air nozzle and their orientation is mostly governed by the consideration of achieving a uniform temperature of oxidation zone. Several investigators have recommended five radial nozzles whereas Groenveld has used a central single nozzle. Susanto has used two nozzles with provision for pyrolysis gas recirculation for minimization of the tar. Bhagwat and Parikh found introduction of a Swirl component beneficial in tar minimization. They have used three nozzles, air entry velocity varying from 3 to 9 m/s and the height from the throat as 100mm, which is in conformity of the findings of Nordenswa. Prasad worked on optimization of the several components in terms of the air entry velocity direction, which was varied from radial to tangential. He concluded that 45° to 60° to the radial direction give better results with reference to tar minimization and overall performance of the gasifier [1]. Channiwala found from his extensive research work that gasifier produces best gas at 45°, best cold gas efficiencies at 60° and minimum tar levels between 60° and 75°. Although in general orientation angle vary between 45° to 75°, best compromise for all the combination is obtained at 60° [1].

B. Criteria for IC engine connected to Gasifier

Acceptable Limits for Tar as shown in table II [5].

TABLE II  
UPPER LIMITS OF BIOMASS GAS TAR AND PARTICULATES [P. BASU]

Application	Particulate (g/Nm <sup>3</sup> )	Tar (g/Nm <sup>3</sup> )
Direct combustion	No limit specified	No limit specified
Syngas production	0.02	0.1
Gas turbine	0.1–120	0.05–5
IC engine	30	50–100
Pipeline transport	-	50–500 for compressor
Fuel cells	-	<1.0

IC engine application requires particulate matter and tar generation quantity at very minute level. This requires intense research with lots of efforts at the design part itself. Different researchers have worked at different levels for best gas quality which are reviewed & basically included in the design part. As gasifier development is defined as art, the variable which affects the performance most are also covered in the design itself. A downdraft gasifier with output capacity of 05 kWe is selected for design for engine application.

III. DESIGN OF VENTURI TYPE GASIFIER

Gasification rate Q<sub>gen</sub> can be found from,

$$\text{Engine efficiency} = \frac{\text{Power Output in kWe}}{\text{Gas generation rate in Nm}^3 \text{ per hour} * \text{Gross CV of gas generated in kJ per kg}}$$

The literature survey suggest gross calorific value of wood is around 4500-5500 kJ/kg. Reed J gives 5506 kJ/ Nm<sup>3</sup> on wet basis & hence for safer side we assumed Gas C. V as 4500 kJ/kg with almost 20% moisture content or less. Engine efficiency assumed as 25% [4].

$$\text{Gas Generation rate } (Q_{gen}) = \frac{\text{Power output}}{\text{Engine efficiency} * \text{Gas CV}} = \frac{5 * 3600}{0.25 * 4500} = 16 \text{ Nm}^3 \text{ per hour}$$

B.mazumdar [6] put diesel replacement factor as 1-% diesel replace which he took 85%.

$$Q_{gen} = 16 \text{ Nm}^3 \text{ per hour} * (1 - 0.15) = 16 * 0.85 = 13.6 \text{ Nm}^3 \text{ per hour}$$

Fuel consumption: 1 kg of wood generates about 2 to 2.5 Nm<sup>3</sup> gas [3].

$$m_1 = \frac{\text{Gas generation rate in Nm}^3 \text{ per hour}}{\text{Dry gas yield in Nm}^3 \text{ per kg}} = \frac{13.6}{2.5} = 5.44 \text{ kg per hour}$$

Air flow rate (Q<sub>a</sub>): For the producer gas generator/engine can be calculated from above gas output result, using following equation by B. Mazumdar:

$$\text{Air flow rate } (Q_a) = \text{Gas flow rate} * \frac{50}{79} = 13.6 * \frac{50}{79} = 8.61 \text{ Nm}^3 \text{ per hour}$$

Hearth load (B<sub>h</sub>): The hearth load B<sub>h</sub> is defined as the amount of producer gas reduced to normal (p, T) conditions (20° C & 1.01325 bar, 1.204 kg/m<sup>3</sup>) divided by the surface area of the



"throat" at the smallest circumference and is usually expressed in  $m^3/cm^2/h$ . Simply it is the ratio of gas flow to throat cross section area. Cross section area of the throat or hearth load is restricted by certain consideration to bring about proper thermal cracking of tar material in the throat region. Experimentally value is in between 0.4 to 0.9 to meet above criteria. Because one kilograms of dry fuel under normal circumstances produces about  $2.5 m^3$  of producer gas, the relation between  $B_g$  and  $B_s$  is given by FAO:  $B_g=2.5 B_s$ . It means  $0.9 m^3$  of gas is produced for each square centimeter of  $c/s$  area at the constriction or throat. Here we are considering maximum hearth load value  $0.9 Nm^3/cm^2-hr$  for further calculations for tar free gas.

Gasifier throat diameter ( $D_{th}$ ): throat area can be calculated by following equation as per [3,4,5]: The entire document should be in Times New Roman or Times font. Type 3 fonts must not be used. Other font types may be used if needed for special purposes.

$$\text{Throat } c/s \text{ area} = \frac{\text{Gas flow}}{\text{Hearth load}} = \frac{13.6}{0.9} = 15.111 cm^3$$

$$\text{Throat } c/s \text{ area} = 15.111 cm^3 = \frac{\pi}{4} d_{th}^2 \Rightarrow d_{th} = 4.39 cms = 43.9 mm \sim 44 mm$$

Alternatively SGR method for throat diameter as per V P Rathod et al. (1): It was found that an SGR (Specific gasification rate) of  $3000 \pm 10 \% Nm^3/h-m^2$  gave the best performance. The axial temperature profiles recorded under different SGR's indicate that constituent processes of gasification take place within about 500 mm from the throat. This suggests that with continuous biomass feed, the hopper height could be limited to about 500 mm which in turn would resolve most of the flow related problems.

$$A_{th} = \frac{\pi}{4} d_{th}^2 = \frac{Q_g}{SGR} \Rightarrow d_{th} = \sqrt{\frac{13.6}{3000} * \frac{4}{\pi}} = 0.07599 m \sim 76 mm$$

For further calculations we consider this as throat diameter as 76 mm seems more reliable as per other literature survey & experts. (Hearth load will be reducing accordingly). As per physical visits to different sites mentioned for different designs & other literature survey suggests 100 mm throat diameter as more realistic for performance & fluidity of material for the mentioned capacity.

Hence it is decided to check performance of gasifier with 44, 76 & 100 mm. i.e. 2, 3 & 4 inch throat.

Particle shape & size ( $d_p$ ): The size and shape of fuel particles determine difficulty of moving and delivering the fuel as well as the behaviour of the fuel once it is in the gasifier. Size and size distribution of the fuel determines the thickness of the gasification zone, the pressure drop through the bed and the minimum and maximum hearth load for satisfactory operation.

Uniform particle size helps overcome some problems. Improving the grate design, as well as added agitation or steering can go a long way to give trouble free gasifier operation and to broaden the range of fuel shapes suitable for gasification. Proper angle of repose between  $45^\circ$  to  $90^\circ$  with smooth hopper walls are always required as per SERI, P.12 [4]. The quarter cylindrical shaped particles of 25 mm size emerge as the best flowing particle in the investigated range of  $D_{shell}/d_p$  of 8 to 20. The ratio of throat diameter to particle size,  $D_{th}/d_p$  may be taken as 5 for free flow condition for the woody biomass material where  $D_{shell}$ =Diameter of shell in mm,  $D_{th}$ =Diameter of throat in mm,  $D_p$ =Diameter of particle in mm [1]. Literature survey suggests that the Quadra cylindrical shape is good for gasification. By cold flow studies,  $(d_{th}/d_p) > 5$  (Limiting conditions)

$$d_{th}/d_p = 5 \Rightarrow d_p = 76/5 = 15.2 mm$$

Fine grained and/or fluffy feedstock may cause flow problems in the bunker section of the gasifier as well as an inadmissible pressure drop over the reduction zone and a high proportion of dust in the gas. Large pressure drops will lead to reduction of the gas load of downdraught equipment, resulting in low temperatures and tar production. Excessively large sizes of particles or pieces give rise to reduce reactivity of the fuel, resulting in startup problems and poor gas quality, and to transport problems through the equipment. A large range in size distribution of the feedstock will generally aggravate the above phenomena. Too large particle sizes can cause gas channeling problems in gasifiers hence avoid it. It must be emphasized that the above empirical design rules are based on experiences with gas producers fuelled by wood blocks varying in size between 3 to 5 x 6 to 8 cm [3]

Hopper diameter ( $D_{shell}$ ): It depends upon the range of  $D_{shell}/d_p$ , which is 20 maximum. [As mentioned above] & as per (1).

$$D_{shell}/d_p = 16 \Rightarrow D_{shell} = 16 * 15.2 = 243.2 \approx 300 mm$$

Height of shell ( $H_{shell}$ ): The minimum value of the ratio of bed height to container diameter may be taken in the range of 1.3 to 1.7 for better flowability in the investigated  $D/d_p$  range (1). On the basis of cold flow studies, the height of shell was taken as 1.4 times the shell diameter,

$$H_{shell} = 1.4 * D_{shell} = 1.4 * 250 = 350 mm (\text{total height if } D_{shell} = 250)$$

$$H_{shell} = 1.4 * D_{shell} = 1.4 * 300 = 420 mm (\text{total height if } D_{shell} = 300)$$

As per V P Rathod et al.[1], dimensions of venturi for up to throat, where taking  $\theta = 30^\circ$ ,

$$\tan \theta = \text{Throat diameter} / H_{shell} \Rightarrow \tan 30^\circ = 76 / H_{shell} = 131.63 mm \approx 150 mm$$

Remaining hopper height is  $350 - 150 = 200$  mm.

Throat orientation for venturi type design: As provision of a tangential velocity component to the air entry is beneficial from point of view of tar minimization. Channiwala found from his extensive research work that gasifier produces best gas at 45°, best cold gas efficiencies at 60° and minimum tar levels between 60° and 75° [1]. Hence it is decided to measure performance of the gasifier with 03 different angles between given range as 45°, 60° & 70°. Adjacent fixture with matching flanges can be made for combination with 03 decided throat diameters i.e 2.,3 & 4 inch.

Nozzle height ( $H_n$ )- Air entry point: Given 1s as reduction time for gas & mean superficial velocity as 0.4 m/s & suggested value as 0.6, the nozzle height in the range of 50 to 100 mm may be selected for improved flowability as per [1].

$$H_n / d_{th} = 0.6 \Rightarrow H_n = 0.6 * 76 = 45.6 \approx 50mm$$

Air entry point is taken to vary in 50-100mm above throat as per the variation in throat fixture.

Nozzle orientation: For venturi type design, Nozzle orientation does not have any significant effect on the flowability of the material under the investigated range of test conditions. Provision of a tangential velocity component to the air entry is beneficial from point of view of tar minimization. In the experimental research gasifier 60° orientations to the radial gave minimum tar without affecting the performance levels [1]. Here taking 60° with vertical.

Numbers of air nozzles & nozzle diameter: Air velocity and nozzle circle diameter are the two factors which governs the number of nozzle required. Not enough and exact data is available to make a definite recommendation. The overall approach should be to ensure a homogeneous oxidation with no dark corners or zones as per FAO. Nozzle ring circle diameter increases implies number of nozzle increases.

The radial air entry velocity of 7 m/s gives the better flowability of the material under the investigated test conditions. Nozzle orientation does not have any significant effect on the flowability of the material under the investigated range of test conditions as per [1]. Here two nozzles are considered. One with 40-50 mm diameter with 30° angle to horizontal and one vertical central nozzle with diameter as 25-35 mm. Height of air nozzle for venturi type design will be adjusted accordingly for 2, 3 & 4 inch diameter with 45°, 60° & 70° cone inlet angle.

Reduction zone configuration: Height of the reduction zone found from literature was at least 20 cm. [3]. The main parameter of reduction zone is its bed height. Effective conversion obtained if the total bed height from the point of air entry is in the range of 250 mm to 400 mm as per different literature survey.

$$\text{Throat length} = d_{th} / 2 = 76 / 2 = 38mm \approx 50mm$$

Taking calculations as per [1] here, so  $L_d = 350$  mm (total length)

Effective length of divergent section zone:  $L_d = 350 - 60 - 50 = 240$  mm ~250 mm.

$$\text{Length of reduction zone} = L_R = d_{th} * 2 = 152mm \approx 200mm$$

Selecting, DR = 200 mm,  $\tan\theta = 14.5^\circ$ . (Throat length will be adjusted according)

Distance between grate and exposed surface of water seal: This distance can be optimized experimentally and should be lies between 200 mm to 250 mm.

Studying the design with different parameters in detail with available results, it seems that venturi type design is suitable for engine application as it is generating tar content with other particulate matter in the nearest range of requirement of engine limit and also within that limit. Drawings as per the dimensions prepared before fabrication for rough idea which are as below.

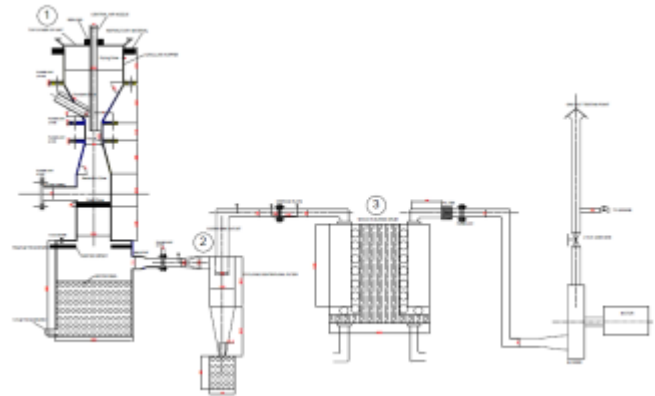


Fig. 1 Venturi type small down draft gasifier with cooling and cleaning arrangement.

Fabrication of both designs with different throat diameters requires air sealed joints & hence it requires much more care at fabrication part. Design drawing represents the rough idea to fabricate the gasifier with different fixtures of throat diameters, angles & with different vertical air nozzle adjustment as per throat length change for optimum performance. There are three variables in this design. One is throat diameter 44, 76 & 100mm, second is inlet throat angle 45°, 60° & 70° & third is nozzle height adjustment 80, 100 & 120mm as per the throat fixture adjustment. Drawings of fixtures are shown below. Hence performance of gasifier to be developed can be check by adopting taguchi design of experimentation method (DOE) for matrix of 3\*3\*3 which gives optimum 9 no of reading to be taken for 27 practicals as per table III below.

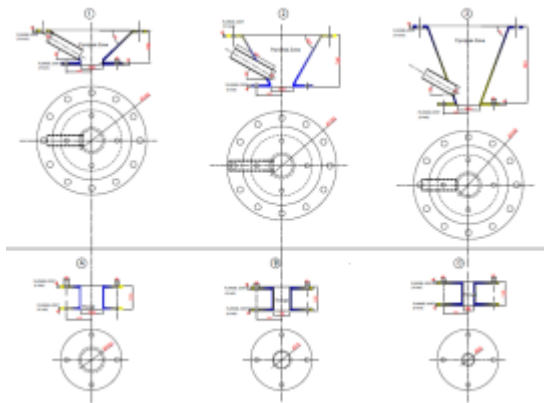


Fig. 1 Design drawings of fixtures of 03 Throat & 03 inlet cone angles

TABLE III  
Layout of L9 Orthogonal Experiments

A: Throat diameter	B: Throat angle	C: Distance of Air injection point from throat
44	45	80
44	60	100
44	70	120
76	45	100
76	60	120
76	70	80
100	45	120
100	60	80
100	70	100

This can be further optimized and compared with existing available results. The final comparison gives optimum gas generation rate with best quality & that can be supplied to engine with developed cooling and cleaning system.

Gas cooling and cleaning system: The temperature of gas coming out of gasifier is normally between 300°-500°C. This gas has to be cooled in order to raise its energy density. Most coolers are gas to air heat exchangers where the cooling is done by free convection of air on the outside surface of heat exchanger. Since the gas also contains moisture and tar, some heat exchangers provide partial scrubbing of gas. Thus ideally the gas going to an internal combustion engine should be cooled to nearly ambient temperature. As per engine requirement, general idea is to put one cyclone separator, one wood filter with suction blower for testing and give it to engine suction as per the maintained air to fuel ratio. Standard design of cyclone is in table 4 as per [2] figure 5 represents the same. Cyclone filters are dry type filter. They are designed according to the rate of gas production and its dust content. The cyclone filters are useful for particle size of 5 µm and grate. Since 60-65% of the producer gas contains particles above 60 µm in size the cyclone filter is an excellent cleaning

device. Small cyclones required for small gasifiers are not available commercially, so they must be custom designed and fabricated. Wood dust in the wood filter requires 1 to 4 mm of dust as per the industrial survey followed by fabric bag filter before engine supply at atmospheric temperature. It was found that before both of this there is water scrubbing arrangement for cooling and cleaning in place of cyclone separator. Water scrubbing is used in some of the designs as per literature survey. One 0.5 HP pump with circular arrangement of holed pipes at a distance almost 200+ mm can be put for cooling and cleaning of outlet gas to grate. Direct approach to grate will be avoided with downwards holes in pipe for washing or scrubbing the gas. Outer 450 mm pipe will seal inside 200 mm pipe and in between 20 to 25 mm pipes will be arranged for spraying water as per the suitable design for cooling and cleaning purpose. (Drawing 3 gives better idea).

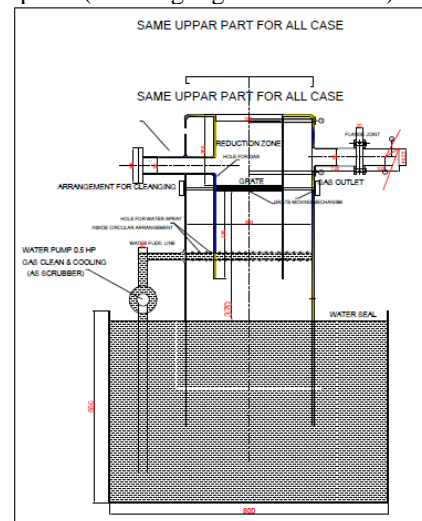


Fig.3 Design drawing of cooling system & water seal.

One filter to protect water carry over will be provided at outlet to gas & water vaporization rate can be calculated for the system. For this arrangement by industrial visit it is concluded that it should be followed by wood dust filter & one fabric bag filter for further cleaning.

With all discussed design concept & prepared design drawings with cooling and cleaning system for gasifier, one 05 kWe gasifier design, fabricate & developed at L D Collage of engineering. 03 inlet throat cones with matching throat diameter fixtures also prepared. As a part of cooling and cleaning system, one tank with spray cooling system followed by cyclone separator & wood filter also fabricated & erected at the place. Figure 4 below shows the developed gasifier.

As a part of performance parameters of gasifier, temperature in different zones of gasifier and pressure drop within the whole system from gas outlet is to be checked. Gas chromatography gives the detailed gas analysis for optimized engine application. With decided air fuel ratio, producer gas can replace the diesel fuel in duel fuel mode or with pure gas engine application



Fig.3 Venturi Type Gasifier at L D Collage O Engineering with cyclone separator, wood filter & blower ready for testing.

As a part of performance parameters of gasifier, temperature in different zones of gasifier and pressure drop within the whole system from gas outlet is to be checked. Gas chromatography gives the detailed gas analysis for optimized engine application. With decided air fuel ratio, producer gas can replace the diesel fuel in duel fuel mode or with pure gas engine application.

#### IV. CONCLUSIONS

Detail study & review of design for small capacity economical downdraft throat type gasifier throws light on following points to use the generated producer gas as alternate fuel in duel fuel arrangement with engines:

The generated gas quality affects much to engine & hence gasifier should ensure as much as tar cracking followed by

design gas generation rate with much lesser tar and other contaminant in the outlet of gas. The main parameters which affect gas quality are throat diameter & angle, nozzle height & orientation with required air flow direction & quantity, reduction zone configuration & arrangement of cooling and cleaning. For optimization, different fixture has been fabricated with different combinations of this parameters & optimum gas quality will be checked for engine application. Optimized gas with lowest tar content & other required parameters will be tested in diesel engine. Diesel engine will be tested in duel fuel mode along with producer gas. Cooling and cleaning arrangement for the gasifier is still found under development and research work is still going on in this area.

#### ACKNOWLEDGMENT

The heading of the Acknowledgment section and the References section must not be numbered.

Causal Productions wishes to acknowledge Michael Shell and other contributors for developing and maintaining the IEEE LaTeX style files which have been used in the preparation of this template. To see the list of contributors, please refer to the top of file IEEETran.cls in the IEEE LaTeX distribution.

#### REFERENCES

- [1] Rathod V P et al., Design development of downdraft venturi type gasifier, ICME03- TH- 15, Dhaka, Proceedings of ICME 2003.
- [2] B. wang, Numerical study of gas- solid flow in cyclone separator, CSIRO, 3rd International conf., Australia, 2003
- [3] Wood gas as engine fuel, Mechanical Wood Products, Branch Forest Industries Division, FAO Forestry Department, ISBN 92-5-102436-7, M-38, , © FAO 1986.
- [4] Reed J, SERI, Handbook of biomass downdraft gasifier engine systems, U S Department of Energy, 7 October, 1988.
- [5] Prabir Basu, Biomass gasification & pyrolysis- Practical Design, Academic Press is an imprint of Elsevier, ISBN 978-0-12-374988-8, 30 Corporate Drive, Suite 400 Burlington, MA 01803, USA, 2010.
- [6] Text Book Of Energy Technology By B.mazumdar,1999, P. 24,25

# Role of RFID in Supply Chain Management of Perishable Products

A Chandrakumar M. Badole, B Dr. Rakesh Jain, C Dr. A.P.S.Rathore, D Dr. Bimal Nepal

**Abstract**--Even after considerable advancement in technology and facilities supply chains of perishable products are suffering from number of complexities and challenges. Inadequate planning, lack of timely information and non-utilization of newer technologies make these chains inefficient as well as non responsive. In particular these chains suffer from lack of control over lead time, its variation and the demand variability in supply chain cycles. This paper discusses how lead time, its variation and demand variability affects the supply chains performance, referring to earlier research. And then suggest that the utilization of newer technology like radio frequency identification (RFID) can be helpful to overcome these problems to certain extent and improves performance of supply chains in many directions and performance measures.

**Index Terms**- Perishable product, shelf life, lead time, demand variability, RFID, supply chain cost.

## 1. INTRODUCTION

Now days the changing global market scenario, high competition, obsolescence of perishable products due to short self life and changing customer demand are among the key challenges faced by perishable supply chains. Inherent uncertainty of supply networks, proliferation of product variety and shortening of product life cycles are among the other factors responsible for revolution in existing perishable supply chains. These supply chains need to compete with growing variety of products, short delivery time, high service level, high quality and lower cost. Perishability imposes an intense pressure on managers to manage the supply of perishable products. Due to perishability, there is an additional cost of disposal of outdated items, and this can also lead to out-of-stock situations, if not managed effectively [1]. The product becomes absolute if it is not getting consumed within its shelf life. In India about 32 % of the fruits and vegetable products go to waste due to insolvency. This waste is large and expensive for both; supply chains as well as to the society.

The management of optimum amount of perishable inventory is the other challenge to satisfy customers. Higher amount of inventory leads to the obsolescence and scarcity leads to loss of reputation of the firm. Product variation increases rapidly due to customer and market fragmentation and specialization [2]. However, the total market volume does not rise as fast as the number of products that are being offered. This leads to a decreased in volume per product type. Therefore satisfying customer is a big problem for retailers. Maintenance of logistics and production systems in order to respond quickly is the biggest challenge faced by

perishable supply chains. There is considerable supply variations due to seasonality of agricultural production, whether conditions and biological nature of agricultural products. This results in input variation and unpredictability.

A perishable product gets spoiled and its useful life reduces if not handled properly during transportation. If this reduced life information of items is not updated, then it may be possible that an outdated item gets delivered to a customer. In such a case, there may be an additional cost of replacement of item and also loss of goodwill of customer. Such spoilage could be reduced simultaneously with automating inventory management, by using RFID technology for product identification, while it moves through the supply chain [3]. RFID system can track the items in real time without product movement, scanning or human involvement. Using active RFID tags it can be possible to update information on it dynamically. RFID system gives a complete visibility of product movement in the supply chain [4]. This may help to make early decisions about inventory control in case there is any interruption in the supply. It partially or completely eliminates time and effort required for counting while loading/unloading the items. This results into reduction of total lead-time for arrival of an order [5].

Two areas that managers focus on are the reduction of the replenishment lead time from suppliers and the variability of this lead time. The normal approximation of lead time demand distribution indicates that both actions reduce inventories for cycle service levels above 50%. The normal approximation also indicates that reducing lead time variability tends to have a greater impact than reducing lead times, especially when lead time variability is large [6]. A larger lead time requires a higher optimal base-stock level [7]. Also when the lead-time is longer, the order rate of the supply chain system become more oscillatory and the bullwhip effect phenomenon become more evidence [8]. This research work is also supported by the work of [9], which gives the same result. To justify this we present the analysis of some of the past research in a greater details. For this we consider the work carried out by Sunil Chopra et al., Xiaojing Wang et al., Jing-Sheng Song and Warburton et al. in more depth in the forthcoming paragraphs.

Over the past decades, there has been a notable amount of contribution by various researchers in the area of perishable product supply chains, specifically oriented towards developing optimal order quantity policies and inventory management policies. But the contributions on the impact of newer technologies like RFID in supply chain of perishable products is negligible and in the

preliminary stage. Yet work of quantifying benefits of RFID is in continuation and now days many researchers are seen inclined towards this task. The brief of the research work in this regards is presented in the following few paragraphs.

The remaining of the paper is organised in following manner. Section 2 highlights the past work in this area. Section 3 discusses the past work of key researchers in details to see the effect of lead time, its variability and demand variation on supply chain performance. In section 4 we present the constructional and working details of RFID system. Section 5 presents the RFID based supply chain architecture and impact of RFID on supply chain. The last section summarises paper with conclusion and future scope for this research.

## 2. REVIEW OF PAST WORK

In the early stage of research work a comprehensive literature review was carried out considering 302 research papers from about 40 reputed international journals, conference proceedings, research thesis and national conference. The supply chain modelling based literature was only considered for reviews that were sourced from science direct, IEEE explore, spingerlink, google scholar etc. The out come of the review is that the research on adoption of new technology like RFID in controlling the supply chains of particularly perishable products is scare and in the preliminary stage. Very few researchers have found worked in this region of supply chain management. Still today the beneficial impacts of this technology have not been quantified mathematically and fruitfully. The researchers are found to have assumed the benefits in certain percentage using different notations.

The research work on supply chain of perishable product have seen started form the mid of nineteenth century and has got inertia at the end of nineteenth century. Credit goes to [10-12], who modelled economic order quantity for perishable inventory. Next [13] developed an inventory model with deterministic but linearly changing demand rate, constant deterioration rate and finite planning horizon. Next,[14] extended Dave and Patel's model to allow for shortages. However, [15] presented an EOQ model considering variables deterioration and power demand pattern. Research on models with deteriorating items, time-varying demand and shortages continues with [16-21]. The common characteristic of all the above papers is that they allow for shortages while unsatisfied demand is completely backlogged [22]. Most of the research work is concentrated on EOQ, replenishment and supply chain costs [23, 24]

In the beginning of the twentieth century the trend has been found changed and the researchers are found inclined towards modelling the pricing strategies. The concept of dynamic pricing was introduced to account for loss of inventory due to obsolescence. In this strategy the researchers suggest that the price of the product be reduced as the self life, quality of product degrades day

by day [25, 26]. This strategy stood helpful in boosting demand for products and better generation of revenue which could have lost in spoilage and loss of goodwill of customers. The work was concentrated on issuing policy of retail store [2] where customers are found to prefer last in first out (LIFO) rather than first in first out (FIFO) policy. This justifies the quality consciousness of the today's customers. In such situations the inventory loss can only be reduced by managing and maintaining the adequate quantity of perishable products and is only possible by having the right time information about inventory at various stages and accordingly replenishing it time to time. The newly emerging RFID technology is found to be a better solution to overcome problem [27].

The researchers like [28] have introduced the RFID with constructional details in their research. Basically the RFID system consists of tag, antenna, sensor and software required for computational work, [29-35] have explored the impacts and benefits of RFID on SC. They have studied the impact of RFID for retailers, assembly automations, grocery retailing, FMCG apply chains, in getting clearances in cargo yards and managing shipment in the yards. However [1] have carried out a research of its own importance where in the benefits of RFID was quantified for perishable SC. Some researchers like [36-39], have carried out the economical assessment of investment and benefits of RFID in supply chains. However, [40] have carried out breakeven analysis for installation of RFID in the business organization. The researchers [41, 42], explored the challenges to implement the RFID technology in an organization. Implementation of RFID is subjected to risks and needs to follow certain protocols, rules and regulations. The risks, privacy and security protocols to be followed if RFID is implemented have been discussed in more comprehensive manner by the researchers [43-46]. Last but not least work carried out by [47-49] adds value to the importance of RFID which can be effectively used in cold chains to increase self life of product and hence to maintain quality for long time.

## 3. EFFECT OF LEAD TIME ON SUPPLY CHAIN OPERATION

Lead time has a pronounced impact on in SC operations and in SC and hence reduction of lead times and their variability is a key element of process improvement. In the past considerable contributions have been found out to study the impact of lead time on SC operations and performance measures. Some of the work reflects its own importance in this context as the researchers have modelled the fact and quantified the effect of lead time. In this section the effect of lead time on SC operation is discussed. Past work of few key researchers in this regard is referred in the chronological order. We refer the work of the researchers namely Jing-Sheng, Song Sunil Chopra et al., and Warburton et al. Xiaojing Wang et al...

### 3.1 Model 1

Here, we refer to the model developed by Jing-Sheng Song [7]. The author has built up model for single-item

inventory where demands form a compound Poisson process and lead times are stochastic. This model studies the impact of lead time, lead time demand and effect of variable lead time on the long run average cost of SC.

The model is (for notations meaning, constraints and assumptions refer the paper).

Expected long-run average cost of system  $i$  under the base-stock policy  $\partial(y)$  is

$$c \lambda E[Z] + C_i(y),$$

Where,

$$C_i(y) = E[h(y - D_i)^+ + p(D_i - y)^+] \\ = (p+h) \sum_{u=0}^{y-1} (\Psi_i u) + p(E[D_i]-y) \dots (1)$$

- $\lambda$  = demand rate
- $c$  = unit order cost
- $Z$  = batch-size random variable, discrete in value
- $h$  = unit holding cost rate
- $p$  = unit penalty cost rate
- $L_i$  = lead time of system  $i$ , a continuous random variable,
- $\partial(y)$  = stationary base-stock policy with base-stock level  $y$
- $D_i$  = lead time demand in system  $i$ , a discrete random variable
- $\Psi_i$  = cumulative distribution function of  $D_i$

The model is solved for minimum average cost considering the effect of lead time, lead time demand and effect of variable lead time.

It is observed that stochastically larger lead-time results in a stochastically larger lead time demand, which in turn results in a higher optimal base-stock level. However, a stochastically larger lead time may not necessarily lead to a higher optimal average cost, because some times the variability effects may dominate. On the other hand, a more variable lead time always leads to a more variable lead time demand, which in turn implies a higher long-run average system cost for any fixed base-stock policy. Consequently, a more variable lead time always leads to a higher optimal average cost. The effect of lead time variability on optimal policies depends on the inventory cost structure: A more variable lead time requires a higher optimal base-stock level if and only if the unit penalty (holding) cost rate is high (low).

### 3.2 Model 2

The second model we consider for analysis is formulated by Sunil Chopra et al.[6]. It says that traditionally, a normal approximation has been used to estimate the relationship between safety stock and demand uncertainty, replenishment lead time, and lead time uncertainty [50]. The researchers [6] have modelled the effect of lead time and lead time variability on reorder point and safety stock as a performance measures for the varying amount of cycle service level provided to the customers. The model was numerically tested for tow criterions namely normal approximations of lead time

demand distribution and the exact demand during lead time. The results contradicts the results of [50] and proves that

1. For cycle service levels above 50% but below a threshold, reducing lead time variability increases the reorder point and safety stock.
2. For cycle service levels above 50% but below a threshold, reducing the lead time variability increases the reorder point and safety stock, whereas reducing the lead time decreases the reorder point and safety stock. However, the results of normal approximation says
  1. For cycle service levels above 50%, reducing lead time variability reduces the reorder point and safety stock.
  2. For cycle service levels above 50%, reducing lead time variability is more effective than reducing lead times because it decreases the safety stock by a larger amount.

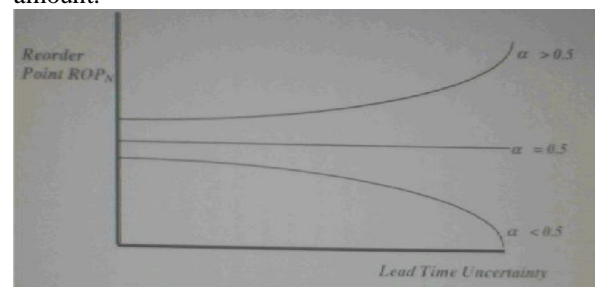


Figure 1: variation of ROP for different CSL for varying LTU in case for normal approximation

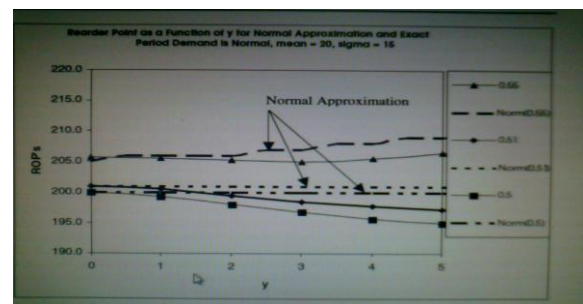


Figure 2: variation of ROP for different CSL for varying LTU in case for normal approximation and the exact distribution.

In short, it is concluded that for cycle service levels above 50% the normal approximation predicts that a manager can reduce safety stocks by decreasing lead time uncertainty. Our analytical results and numerical experiments, however, indicate that for cycle service levels between 50% and a threshold, the prescription of the normal approximation are flawed and decreasing the lead time uncertainty, in fact, increases the required safety stock. In this range of cycle service levels, a manager who wants to decrease inventories should focus on decreasing lead times rather than lead time variability. This contradicts the conclusion drawn using the normal approximation.

3.3 Model 3

Next we consider the model constructed by Warburton et al. [8]. The researchers have studied the impact of bullwhip effect arises due to larger lead times on the inventory level in the supply chain. The analysis starts with the simple supply chain equations. Retailers attempt to minimize their inventory while maintaining sufficient on hand to guard against fluctuations in demand. The inventory,  $I(t)$ , at any time (t) is depleted by the demand rate,  $D(t)$ , and increased by the receiving rate,  $R(t)$ , so the inventory balance equation is:

$$\frac{dI}{dT} = R(t) - D(t) \dots\dots\dots (2) \text{ And}$$

$$R(t) = O(t - \tau) \dots\dots\dots(3)$$

For the demand term, a step function surge in demand is analyzed, which is a rich source of insight when seeking an understanding of the trade-offs involved in tuning an ordering policy. Also, since the equations are linear, any arbitrary demand can be built from a suitable linear combination of step functions. The lead-time,  $\tau$  is the time from the issue of orders until the receipt of the goods from the supplier. Thus, the receipts are equal to the orders placed at a previous time. The second equation merely expresses the fact that the receiving rate is equal to the order rate with lead time  $\tau$ . It is usually suggested that some kind of smoothing should be applied to the demand data. Otherwise, excessive fluctuations occur resulting in increased production costs. Exponential smoothing is easy to implement and relatively accurate for short-term forecasts which is suitable for perishable products. The tunable parameter,  $Ta$ , controls the amount of smoothing to be applied to the raw demand. The smoothed demand is given by equation

$$Ds(t + \hat{c}t) = Ds + \frac{(AD(t) - Ds(t))\hat{c}t}{Ta} \dots\dots\dots(4) \text{ And the}$$

smooth demand contribution to the order rate is then

$$Od(t) = Do + d\{1 - \exp(-t/Ta)\} \dots\dots(5)$$

The inventory replenishment goal in the ordering policy is to bring the actual inventory towards the desired inventory:

$$Oi(t) = \frac{Io - I(t)}{Ti} \dots\dots\dots(6)$$

$Io$  represents the desired inventory. This policy has the advantage that it replaces deficits due to a surge in demand, and the tunable parameter,  $Ti$ , acknowledges that the deficit recovery should be spread out over time  $T$ . The order rate and inventory equations make up the system to be solved. Equations (2) and (3) can be solved exactly in terms of the Lambert W function. [51]

The result is that the entire, exact solution for the inventory and orders is:

$$I(t) = Io - dt \text{ for } t \leq \tau \text{ and } \dots\dots\dots(7)$$

$$I(t) = ID - dT + A \exp[Wt / \tau] \dots\dots\dots(8) \text{ Where,}$$

$$W = W(-\tau / T) \dots\dots\dots(9) \text{ the Lambert W function.}$$

The simulation results of the solutions can be seen in Figure 3, where it is clear that the response of the inventory depends sensitively to the ratio of the replenishment delay,  $\tau$ , to the inventory deficit parameter,  $Ti$ . Larger replenishment delays increase the divergence of the inventory response (large overshoots). However, the parameter,  $Ti$ , can be tuned so that the inventory returns exactly to the desired level without overshoot. The inventory response for the critical value,  $Ti^*$  is also shown in the figure 3.

Also according to the author, the bullwhip effect (BW) is the ratio of retailer order rate to the demand rate. Therefore,

$$BW = \frac{O(\tau)}{d} = \frac{Id - I(\tau)}{Td} = \frac{\tau}{T} \dots\dots\dots(10)$$

Thus the results of figure 3 equally holds good for equation (10)

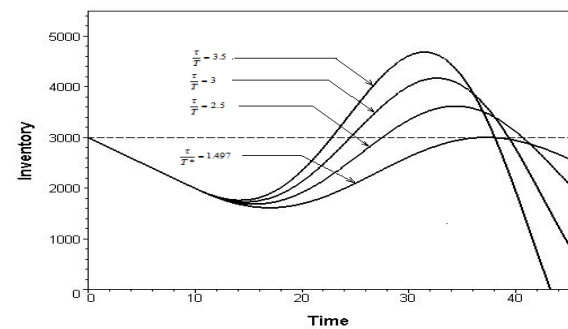


Figure 3: Effect of lead time variation on the inventory level.

From analysis of model 3, it is conferred that the lead time and its variation has definitely affects the inventory level in the supply chain. Higher lead time results in boosting the bullwhip effect, higher inventory and higher costs.

3.4 Model 4

In the last we consider the model of Xiaojing Wang et al.[9]. The simulation results of this model equally supports the results of above model and infer that the order rate

is influence by the lead time, and the longer the lead-time the more oscillatory the order rate. Furthermore, higher the lead time boosts the bullwhip effect. In addition, the



analysis of the impact of lead-time on the total inventory level and net inventory level reveals that reducing the lead-time can attenuate the risk of stock out and smooth the inventory level.

From the analysis of these four models, it becomes evident that in supply chain the lead time is the important factor for decision making. The managers must try to control lead time, lead time variations and lead time demand variations to reduce inventory in hand sufficient to satisfy customer demand without shortage. This is equally important for the supply management of perishable product to reduce loss due spoilage and to provide fresh products to the customer with optimum cycle service level. This objective can be definitely achieved by incorporating newer technology like RFID in the supply system.

#### 4. THE RFID SYSTEM

Supply chain management objective is to increase the long-term performance of individual companies and the overall supply chain by maximizing customer value and minimizing costs. Not all companies achieve these goals with the same strategy. A supply chain is either agile or lean and given this, a different approach to increase the efficiency and effectiveness is adopted. Companies such as Wal-Mart and Dell have gained efficiencies by having a clear understanding and a tight commitment to deliver customer value by maximizing not only the value provided by their companies but also aligning their partner's interest to create unique supply chains. Information systems are the backbone of every supply chain and they are based on automatic data acquisition techniques to meet the goal of collecting information. RFID technology with unique characteristics can make it suitable to enhance data collection processes along the supply chain.

##### 4.1 What is RFID?

Radio Frequency Identification (RFID) is an automatic identification and data capture technology which is composed of three elements: a tag formed by a chip connected with an antenna; a reader that emits radio signals and receives in return answers from tags, and finally a middleware that bridges RFID hardware and enterprise applications. RFID technologies with the appropriate IT infrastructure help both major distributors and manufacturers, as well as other logistics operations, such as the health-care system, defence industries, and others, dealing with complex, global supply chains in which products and product shipments must be traced and identified in a non-contact, wireless fashion using a computer network, because of cost, or security, or safety, or because parts are subject to corrosion, or food/medicine is subject to quality degradation, or other reasons [52]. All of these requirements point to an automated wireless, readable sensory-based identification method, and network, that offers more functionalities and is significantly valuable than the existing methods.

RFID has the potential to change the way we do business all around the world [53]. It is a huge challenge, not just because of the sophisticated sensor-network technology, but also because of the vast systems integration and coordination. Radio frequency identification (RFID) technologies with the appropriate IT infrastructure help major distributors and manufacturers, as well as other logistics operations, such as the health-care system, defence industries, and others, dealing with complex, global supply chains in which products and product shipments must be traced and identified in a noncontact, wireless fashion using a computer network [54-56].

RFID technology can increase a company's efficiency and provide other benefits to both companies and consumers; however, RFID, like any newly implemented technology, presents management with issues of new system threats and decisions about incorporating adequate controls over the new technology [57]. There are many valid reasons for carrying out research and deployed in industry the wireless, computer networked, sensory-based part identification methods, tools and technologies. The application fields and opportunities are vast. The key driver is that even in chaotic, largely distributed, more stochastic than deterministic business environments, adaptive organizations and enterprises must react to demands quickly, else a competitor will take the business. Therefore, they must reduce waste and improve efficiency at all fronts. The most important aspect of this strategy is to know exactly what parts they have in stock, exactly where these parts are, and in what condition/state of assembly, or preparedness [27].

Furthermore, major distributors dealing with complex, global supply chains must be able to trace their shipments in detail, either because of cost, security, safety, quality degradation (as it is in the case of temperature, humidity, and/or shock sensitive components or drugs), and other reasons.

While the improved information accuracy through RFID deployment will allow companies to substantially reduce out-of-stocks and back orders, they are also likely to find themselves with higher overall average inventory. This suggests a remarkable improvement opportunity, namely that companies can potentially reduce reorder quantities and target inventory levels without hurting customer service levels. This opportunity to reduce inventory and, at the same time, improve customer service levels can be applied not only across multiple tiers in the supply chain, but also within a single store (backroom and on-shelf replenishment). For example, accurately knowing the on-shelf inventory of products will enable replenishment of the shelf stock in an on-demand manner, thereby allowing the company to reduce the overall average store inventory without sacrificing product availability to customers.

##### 4.2 Constructional Details:

This section describes the basic components of an RFID system. Note that the exact configuration of a particular deployment depends on the vendor, system integrator, and the application. An RFID system includes: [27]

- Transponders (Tags) that allow items to be identified.
- Antennas and Readers/writers that allow tags to be interrogated and to respond
- Software that controls the RFID equipment manages the data and interfaces with enterprise applications.

**TRANSPONDERS**

Transponders are the distinguishing feature of an RFID system. Transponders are the ‘labels’ that are attached to objects to be identified. Due to their wide-spread use in supply chain operations, they are commonly known as tags. These microchip-based tags with a tiny antenna attached to them can hold up to 10 Kbits of data. The data stored can include product identification, expiration, warranty, handling and storage instructions, and service history. These tags use radio waves to communicate with the readers.

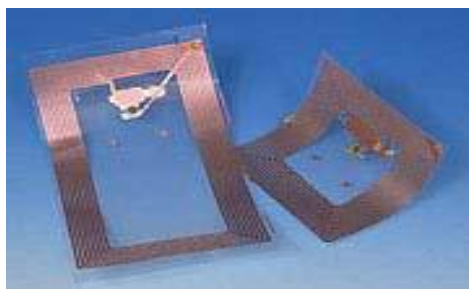
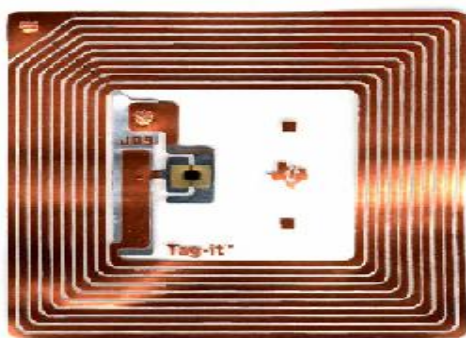


Figure:1 (a) Active Tag

To produce radio waves, tags require some source of energy to power its electronics. Active tags (Figure 1 a) use a tiny battery, a microchip, and a tiny antenna built into them. Passive tags, such as the one shown in (Figure 1 b), are relatively inexpensive, do not contain a battery within them, instead they draw power from the reader’s radio signals that induce a current in the



**ANTENNAS**

Antennas are used in both the tags and the reader. The size could vary from under a square centimetre to several. square meters. UHF reader antennas can be

Figure: 1 (b) Passive Tag

tag’s antenna using either inductive coupling or electromagnetic capture. This power is used both for chip operation and for communication. The operating frequency of radio waves employed also varies. Low-frequency RFID tags operate at 125 to 134 kHz, however high-frequency systems uses 13.56 MHz. Frequencies of 866 to 960 MHz are used in ultra-high-frequency systems, while microwave systems operate at 2.4 to 5.8 GHz. A comparison between the active and passive tags is shown in the following tables 1 and 2 [8].

Table 1. Characteristics of Active and Passive Tags

Tag Approximate Frequency	General Tag Type	Approximate			
		Range	Transmission Rates	Power Consumption	Cost
Low High	Passive	< 1.0 m	1 – 2 kb/s	20 μ W	\$0.2 - \$1.0
		1.5 m	10 – 20 kb/s	200 μ W	\$1.0 - \$10
Ultra High	Active	10 – 30 m	40 – 120 kb/s	0.25 – 1.0 W	\$10 - \$30

Table 2. A Comparison of High and Low Frequency Tags

Tag Frequency	Relative Range	Transmission Rates	Power Consumption	Relative Cost	Environmental Susceptibility
Low	Shorter	Lower	Lower	Lower	Lower
High	Longer	Higher	Higher	Higher	Higher

classified as circular-polarized or linear-polarized antenna. The former emit and receive radio waves from all directions, while the later work best in one particular direction. (Figure 2). Therefore circular-

polarized antennas are less sensitive to transmitter-receiver orientation and work better ‘around corners’. However, the operating range of a linear-polarized antenna is more than that of a circular-polarized antenna.

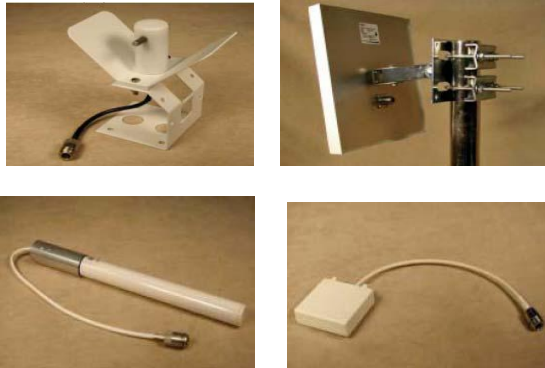


Figure 2. Types of Antenna from: left to right Corner, Flat panel, Omni, and Dual flat panel

The operating range of an antenna also depends upon its ability to focus radio waves. Higher gain antennas can work at a longer range than the lower gain counterparts. However, there is always a trade-off between gain and coverage area. When a tag communicates with an antenna, the radio frequency portion of the circuit between the tag and the antenna is called the air interface. This radio communication takes place under a certain set of rules called air interference protocol. Propriety protocols may cause interoperability problems with equipment from different vendors.

**READERS**

Readers read or interrogate the tags. In reading, the signal is sent out continually by the (active) tag whereas in interrogation, the reader sends a signal to the tag and listens. To read passive tags, the reader sends radio waves to them, which energize them and they start broadcasting their data. The reader reads all the tags within its read range in a quick succession. This automatic process reduces read times. If more than one tag is present within the range of a reader, various techniques are available to read them all sequentially. These techniques, grouped under the name of simulation identify individual tags by allowing only the tags with a specific serial numbers to respond. The scheme where the reader controls the response timing of the tags is known as reader talks first method. Conversely, the scheme where tags start beaming their data as soon as they are energized by the reader is known as tag talks first method. The former method is more accurate but is slower compared to the latter.



Figure 3: (a) Handheld Reader



(b) Large Frame Reader

Readers work at different frequencies, from a low of about 100 KHz to a high of about 5.8 GHz. The Tris S2000 reader used in our laboratory experiments works at 134.2 kHz, which is at the low end. Read ranges for lower frequency systems are smaller but the systems are less susceptible to performance degradation in presence of metal. or water in the environment.

**SOFTWARE**

Software is the glue that integrates an RFID system. Again it depends upon the industry context, but usually a front end component manages the readers and the antennas and a middleware component routes this information to servers that run the backbone database applications. For example, in a manufacturing context, the enterprise software will need to be made aware of RFID at various levels depending on how far downstream into manufacturing and out into the supply chain RFID is implemented. The RFID Journal categorizes middleware technologies into three levels:

- Software applications which solve connectivity problems and monitoring in specific vertical industries.
- Application managers that connect disparate applications within an enterprise, and
- Device brokers that connect applications to devices like shop-floor machines and RFID readers.

The Auto-ID Centre at Massachusetts Institute of Technology developed a software program named 'Savant' to manage the enormous amount of data expected to be generated by RFID readers. In a typical manufacturing scenario, for example, readers will be picking up a continuous stream of tag data, which might contain errors such as duplicate reads and phantom reads. The job of a savant is to filter and manage this data and forward only clean data in order to avoid overwhelming enterprise applications.

## 5. BENEFITS OF RFID IN INVENTORY CONTROL OF PERISHABLE PRODUCTS:

Effective perishable inventory management depends upon consolidating, integrating, and analyzing data collected from many sources such as, distribution centres and warehouses. Conventional tracking systems require manual intervention, which is labour intensive, time consuming, and error-prone. On the other hand, the use of RFID technology has significant advantages over the conventional methods; these are discussed below and are depicted in figure 5 below [58]

### 5.1 Perishable inventory control:

A perishable product has limited useful life and if it is not handled properly while transporting, it may get spoiled and its useful life reduces. If this reduced life information of items is not updated, then it may be possible that an outdated item gets delivered to a customer. In such a case, there may be an additional cost of replacement of item and also loss of goodwill of customer. Such spoilage could be reduced simultaneously with automating inventory management, by using RFID technology for product identification, while it moves through the supply chain. RFID system can track the items in real time without product movement, scanning or human involvement. Using active RFID tags it can be possible to update information on it dynamically.

### 5.2 Reduced Bullwhip effect:

Exaggeration of demand in upward direction in a supply chain network is termed as Bullwhip effect. Due to tracking limitations of conventional systems it may not be possible to get accurate information on actual sales of items; that will amplify the magnitude of the bullwhip effect. If RFID systems are used for information collection, accurate and real time information on product sale can be captured and used for decision making. This will definitely help to reduce overall bullwhip effect. Reducing bullwhip effect would benefit industries where instances of supply-demand imbalances have high costs attached to it.

### 5.3 Lead-time Reduction:

Conventional systems limit tracking of items while being transported. RFID Systems gives a total visibility of product movement in the supply chain. This may help to make early decisions about inventory control in case there is any interruption in the supply. It partially

or completely eliminates time and effort required for counting while loading/unloading the items. This results into reduction of total lead-time for arrival of an order. Pharmaceutical industry, perishable product industry could use RFID systems for reducing lead-times that will help to increase total useful shelf life of items.

### 5.4 Issuing policies

(FIFO/LIFO): RFID systems give exact count and location of items. This will help to follow a certain issuing policies for items as per the requirements. E.g.: First-in-first-out (FIFO) policy for items such as, vegetables, bread that retailer prefers or last-in-first-out (LIFO) for blood banks and other perishable product supply chains that customer prefers.

### 5.5 Asset Visibility

In a replenishment-based system, whenever the total inventory at a warehouse or distribution centre drops below a certain level, the RFID enabled system could place an automatic order. RFID-tagged products will allow stores to track the location and count of inventories in real time. This will better monitor demand for certain products and place orders to prevent an out-of-stock situation. The high levels of inventory monitoring obtained using RFID can particularly benefit FMCG industries

### 5.6 Point of Sale Data

On the retailing side, RFID technology at the point-of-sale (PoS) can be used to monitor demand trends or to build a probabilistic pattern of demand. This application could be useful for apparel industry or products exhibiting high levels of dynamism in trends.

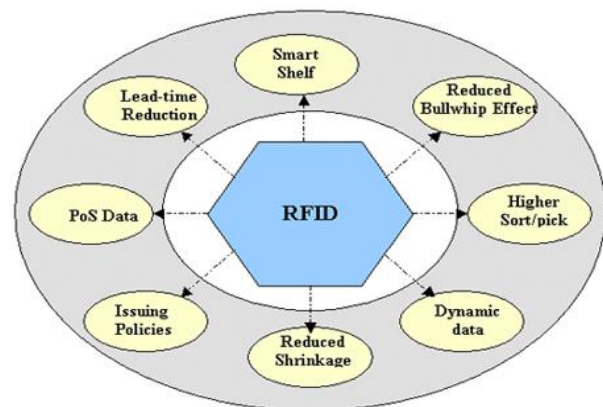


Figure5: Potential benefits of RFID system.

### 5.7 Improves Sort/Pick Rate

In a warehouse, sorting/picking activity is more time consuming and subjected to errors. For example, for issuing of items from a store, a person has to find out whether the item is available in store or not by physically moving to the location. Items issued should be kept in a particular position (bar-code upward) in a pallet for scanning/billing purposes. RFID systems ease

the sorting and picking operations, as it captures real-time, accurate information about product availability in host computer database without physical movement. RFID tags are read via radio frequencies therefore it is not mandatory to place the items in a particular position to read it. This could be helpful for effective warehouse management.

#### 5.8 Reduced inventory shrinkage

As items are continuously monitored, Inventory shrinkages including thefts, misplacement of items and shrinkage due to obsolescence can be avoided using RFID technology.

### 6. CHALLENGES TO RFID ADOPTION

It is proved by several researchers [30] that the RFID can lead organisation to better future. Several benefits of it are highlighted by the researchers in this area. Because of its potential to revolutionize global supply chain management (SCM) systems, RFID was recently the cause of much optimism. Wal-Mart mandated its top 100 suppliers to begin using RFID on 1 January 2005; this day was viewed as a watershed day in the industry. However, the expected rapid industry adoption of RFID has not taken place [41]. In the early stage RFID adoption is a very difficult task. Highly technological skill, mentality to change and large amount of investment is needed to adopt the technology. Other than this the technical problems are always in the path of its adoption. Following paragraph highlights the challenges to RFID adoption.

#### a. Technology challenges

- Material effects on antenna power pattern
- Tag antenna orientation affects radio wave reception
- Collision caused by simultaneous radio transmission

#### b. Standard challenges

- Lack of a unified RFID standard
- Lack of consistent UHF spectrum allocation for RFID

#### c. Patent challenges

#### d. Cost challenges

- Manufacturing costs
- Customization costs

#### e. Infrastructure challenges

#### f. Return on Investment challenges

#### g. Barcode to RFID migration challenges

### 7. CONCLUSION AND FUTURE SCOPE

It is observed from the above analysis that supply chains of perishable products are facing large problems for their survival. Major challenge to them is supply of a fresh quality product with better cycle service level and at low cost by reducing the waste due to spoilage. Lead time, and its variability has a pronounced impact on supply chain performance which can be controlled on supply chain performance which can be controlled using newer technologies like RFID. RFID in spite of

big challenges has proven to be the best solution for inventory control to the companies that have incorporated it. Till date the research work on benefits of RFID is in the preliminary stages. In future there is large scope to develop mathematical model to quantify the benefits of RFID and implement it cost effectively to the supply chains. Our future research is aimed towards the development of framework and model for quantifying the benefits of RFID in supply chains of perishable products.

#### References:

1. Rajurkar S. W., Jain R., 2010. Food supply chain management: review, classification and analysis of literature, International Journal of Integrated Supply Management 6(1) 33-72.
2. David C. Twist, "The impact of radio frequency identification on supply chain facilities", Journal of Facilities Management, Vol.3 NO.3 pp. 226-239
3. Wei Zhou, (2009). "RFID and item-level information visibility", European Journal of Operational Research vol.198, pp. 252-258
4. Peter Jones, Colin Clarke-Hill, David Hillier and Daphne Comfort, (2005), "The benefits, challenges and impacts of radio frequency identification technology (RFID) for retailers in the UK, Marketing Intelligence & Planning Vol. 23 No. 4, 2005 pp. 395-402.
5. Chopra S., Reinhardt G., Dada M., 2004. The Effect of Lead Time Uncertainty on Safety Stocks, Decision Sciences 35(1).
6. Jing-Sheng Song, 1994, "The Effect of Lead time Uncertainty in a Simple Stochastic Inventory Model," Management Science, Vol. 40, No. 5, pp. 603-613
7. Xiaojing Wang, Zhixue Liu, Changzheng Zheng, Chunguang Quan, 2008, "The Impact of Lead-time on Bullwhip Effect in Supply Chain," IEEE Transactions.
8. Roger D. H. Warburton, 2004, "An Analytical Investigation of the Bullwhip Effect", Production And Operations Management Vol. 13, No. 2, pp. 150-160.
9. P.M. Ghare, G.F. Shrader, A Model for exponentially decaying inventories, Journal of Industrial Engineering 14 (1963) 238-243.
10. R.P. Covert, G.C. Philip, An EOQ model for items with Weibull distribution deterioration, Am. Inst. Ind. Eng. Trans. 5 (1973) 323-326.
11. Y.K. Shah, An order-level lot-size inventory for deteriorating items, Am. Inst. Eng. Trans. 9 (1977) 108-112.
12. U. Dave, L.K. Patel,  $\delta T$ ;  $S_i P$  policy inventory model for deteriorating items with time proportional demand, J. Oper. Res. Soc. 32 (1984) 1013-1019.
13. R.S. Sachan, On  $\delta T$ ;  $S_i P$  inventory policy model for deteriorating items with time proportional demand, Journal of Operation Research Society 35 (1984) 1013-1019.
14. T.K. Datta, A.K. Pal, Order level inventory system with power demand pattern for items with variable rate of deterioration, Indian J. Pure Appl. Math. 19 (1988) 1043-1053.
15. A. Goswami, K.S. Gludhuri, An EOQ model for deteriorating items with shortages and a linear trend in demand, J. Oper. Res. Soc. 42 (1991) 1105-1110.
16. L. Benkherouf, On an inventory model with deteriorating items and decreasing time-varying demand and shortages, Eur. J. Oper. Res. 86 (1995) 293-299.
17. M. Hariga, Optimal EOQ models for deteriorating items with time-varying demand, J. Oper. Res. Soc. 47 (1996) 1228-1246.
18. T. Chakrabarti, K.S. Chaudhuri, An EOQ model for deteriorating items with a linear trend in demand and

- shortages in all cycles, *Int. J. Prod. Econ.* 49 (1997) 205–213.
19. M. Hariga, A. Alyan, A lot sizing heuristic for deteriorating items with shortages in growing and declining markets, *Comput. Oper. Res.* 24 (1997) 1075–1083.
  20. J.T. Teng, M.C. Chen, H.L. Yang, Y.J. Wang, Deterministic lot size inventory models with shortages and deterioration for fluctuating demand, *Oper. Res. Lett.* 24 (1999) 65–72.
  21. K. Skouri, S. Papachristos, 2002. “A continuous review inventory model, with deteriorating items, time-varying demand, linear replenishment cost, partially time-varying backlogging,” *Applied Mathematical Modelling* 26 , 603–617
  22. A.K. Bhunia, M. Maiti, An inventory model of deteriorating items with lot-size dependent replenishment cost and a linear trend in demand, *Appl. Math. Model.* 23 (1999) 301–308.
  23. S.K. Goyal, D. Morrin, F. Nebebe, The finite horizon trended inventory replenishment problem with shortages, *J. Oper. Res. Soc.* 43(1992) 1173–1178.
  24. Chung-Yuan Dye, Tsu-Pang Hsieh, Liang-Yuh Ouyang, (2007), “Determining optimal. selling price and lot size with a varying rate of deterioration and exponential partial. Backlogging”, *European Journal of Operational Research*, vol. 181, pp. 668–678.
  25. Ek Peng Chew, Chulung Lee, Rujing Liu, (2009), “Joint inventory allocation and pricing decisions for perishable products”, *Int. J. Production Economics*, vol.120, pp. 139–150.
  26. Alp Ustundag, Mehmet Serdar Kılınc, Emre Cevikcan, (2010), “Fuzzy rule-based system for the economic analysis of RFID investments,” *Expert Systems with Applications*, vol. 37 pp. 5300–5306.
  27. Paul G. Ranky, (2006), “An introduction to radio frequency identification (RFID) methods and solution” *Assembly Automation*, vol. 26, no.1 pp.28–33.
  28. C.M. Roberts (2006), “Radio frequency identification (RFID), computers & security, vol. 25, pp. 18–26.
  29. Peter Jones, Colin Clarke-Hill, David Hillier and Daphne Comfort, (2005), “The benefits, challenges and impacts of radio frequency identification technology (RFID) for retailers in the UK, *Marketing Intelligence & Planning* vol. 23 no. 4, 2005 pp. 395-402.
  30. Eleonora Bottani, Antonio Rizzi (2008), “Economical assessment of the impact of RFID technology and EPC system on the fast-moving consumer goods supply chain”, *Int. J. Production Economics*, vol. 112, pp. 548–569.
  31. Eleonora Bottani, Roberto Montanari, Andrea Volpi, (2010), “The impact of RFID and EPC network on the bullwhip effect in the Italian FMCG supply chain”, *Int. J. Production Economics*, vol. 124, pp. 426–432.
  32. Edmund Prater, Gregory V. Frazier, Pedro M. Reyes, (2005), “Future impacts of RFID on e-supply chains in grocery retailing”, *Supply Chain Management: An International Journal*, vol. 10/2, pp.134–142.
  33. Kazim Sari (2010), “Exploring the impacts of radio frequency identification (RFID) technology on supply chain performance”, *European Journal. of Operational Research*, vol.207, pp.174–183.
  34. James Jungbae Roh, Anand Kunnathur, Monideepa Tarafdar, (2009), “Classification of RFID adoption: An expected benefits approach”, *Information & Management*, Vol. 46 pp. 357–363.
  35. Shiou-Fen Tzeng, Wun-Hwa Chen, Fan-Yun Pai, (2008), “Evaluating the business value of RFID: Evidence from five case studies”, *Int. J. Production Economics*, vol.112, pp.601–613.
  36. In Lee, Byoung-Chan Lee (2010), “An investment evaluation of supply chain RFID technologies: A normative modeling approach”, *Int. J. Production Economics* vol.125.
  37. Jindae Kim, Kaizhi Tang, Soundar Kumara, Shang-Tae Yee, Jeffrey Tew (2008), “Value analysis of location-enabled radio-frequency identification information on delivery chain performance”, *Int. J. Production Economics*, vol. 112, pp. 403–415.
  38. Simon Veronneau, Jacques Roy (2009), “RFID benefits, costs, and possibilities: The economical. analysis of RFID deployment in a cruise corporation global service supply chain”, *Int. J. Production Economics*, vol.122, pp. 692–702.
  39. A.G. de Kok, K.H. van Donselaar, T. van Woensel, (2008), “A break-even analysis of RFID technology for inventory sensitive to shrinkage”, *Int. J. Production Economics*, Vol. 112, pp. 521–531.
  40. N.C. Wu, M.A. Nystrom, T.R. Lin, H.C. Yu, (2005), “Challenges to global. RFID adoption”, *Technovation*.
  41. Kirk A. Patterson a, Curtis M. Grimm b, Thomas M. Corsi c, (2003), “Adopting new technologies for supply chain management”, *Transportation Research Part E*, vol. 39, pp. 95–121.
  42. Roberto Di Pietro, Refik Molva, (2010), optimal probabilistic solution for information confinement, privacy, and security in RFID systems”, *Journal. of Network and Computer Applications*,
  43. Frederic Thiesse, (2007), “RFID, privacy and the perception of risk: A strategic framework”, *Journal of Strategic Information Systems*, vol.16, pp. 214–232.
  44. Ton van Deursen, Sasa Radomirovi, (2009), “Security of RFID Protocols – A Case Study”, *Electronic Notes in Theoretical. Computer Science*, vol. 244, pp. 41–52.
  45. Eun-Kyung Ryu, Tsuyoshi Takagi (2009), “A hybrid approach for privacy-preserving RFID tags”, *Computer Standards & Interfaces*, vol. 31, pp.812–815.
  46. Estrada-Flores S, Tanner DJ, Amos ND. (2002) Cold chain management during transport of perishable products. *Food Australia*, 54 (7): 268-270.
  47. Emond JP. 2007. Quantifying RFID’s cold chain benefits. *RFID Journal LIVE! 2007 Conference*.
  48. Jedermann R, Lang W. 2007. Semi-passive RFID and beyond: steps towards automated quality tracing in the food chain. *Int. J. Radio Frequency Identification Technology and Applications*, 1(3); 247-259.
  49. Eppen, G. D., & Martin, R. K. (1988). Determining safety stock in the presence of stochastic lead time and demand. *Management Science*, 34(11), 1380–1390.
  50. Corless, R. M., G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey and D. E. Knuth, 1996, On the Lambert W function, *Advances in Computational Mathematics* 5, 329-359.
  51. Wilke, P. and Braunl, T. (2001), “Flexible wireless communication network for mobile robot agents”, *Industrial Robot*, Vol. 28 No. 3, pp. 220-33.
  52. Ashley, S. (2004), “Penny-wise smart labels”, *Scientific American*, vol. 291 no. 2, pp. 30-1.
  53. Brzozowski, C. (2004), “Tags, tickets & labels: new technologies emerge”, *Printing News*, vol. 153 No.3.
  54. Glidden, R. and Bockorick, C. et al. (2004), “Design of ultralow- cost UHF RFID tags for supply chain applications”, *IEEE Communications Magazine*, vol. 42 No. 8, pp. 140-151.
  55. Knights, P.F., Henderson, E. and Daneshmend, L.K. (2004), “Drawpoint control using radio frequency identification systems”, *CIM Bulletin*, vol. 89 no. 1003, pp. 53-58.
  56. Leslee N. Higgins and Tim Cairney (2006). “RFID Opportunities and Risks,” *The Journal of Corporate Accounting & Finance*, pp.51-57.
  57. Rfid Application In Inventory Control, The decision makers crafting intelligent solutions, issue no: 03/06/1

# Graph Theory Analysis of Integrated Manufacturing System

A. P. Deb    B. V. Singh

**Abstract**— In this paper, a detailed investigation and comparison of interaction procedures among various subsystems of an Integrated Manufacturing Systems (IMS) through critical observation and analysis has been carried out. The effect of such interactions among various subsystems on the integrated manufacturing system has also been analyzed. A graph theoretic model, an analytical method is used for such purpose. In present work, matrix algebra and permanent function have been used to obtain the permanent multinomial of IMS which forms the basis for obtaining the sub-graphs of the IMS. In the present study the interaction loops among various subsystems and various important aspects towards the fulfillment of different functions have been investigated. The present analysis has shown different nature of activities of various subsystems, the utility of various subsystems, the pattern of information-exchange among various subsystems and the chronological development of exchange of information among various subsystems

**Index Terms**— Integrated manufacturing system, Systems analysis, Graph theoretic approach, Interactions, sub-graphs.

## I. INTRODUCTION

Integrated manufacturing system (IMS) is considered to be a combination of various manufacturing subsystems (e.g., input subsystem (IS), management and control subsystem (MCS), manufacturing subsystem (MS), support and information processing subsystem (SIPS) output subsystem (OS) and new design subsystem (NDS)). Essentially, input subsystem provides raw material and consists of suppliers, vendors etc.; management and control subsystem coordinates and controls the efforts by different subsystems; manufacturing subsystem adds real value in the system; support and information processing subsystem provides means for coordination through sharing of information; output subsystem delivers finished products to market through various outlets and

process design subsystem brings in new product ideas and addresses related issues.

System modeling is a useful tool for analysis of complex systems such as an integrated manufacturing system. Many approaches are there for such analysis like the cause and effect diagrams, event tree diagrams while none of them is all-embracing (Mason-Jones et al, 1998). Some researchers have developed systems models based on graph theory and permanent function to analyze various systems. Graph theory is a well developed field of mathematics (Deo, 2000) while the permanent function is also well researched mathematical operator (Jurkat and Ryser, 1966; Minc, 1966). Venktasami and Agrawal (1996; 1997) have used graph theoretic model for system and structural analysis of an automobile vehicle. Rao and Gandhi (2001; 2002) have used graph theory for decision making in manufacturing engineering domain but they have not discussed about system and structural analysis. Structural modeling integrative analysis of total manufacturing system using graph theoretic approach has been reported by Singh, Agrawal and Deb (2008, 2010) for a manufacturing system consisting of five subsystems. However, detailed interactions among manufacturing design subsystems and processes and their effect on the integrated manufacturing system have not been reported. In the present paper, integrated manufacturing system has been analyzed through critical observation and analysis for continuous improvement in achieving various functions/goals of the integrated manufacturing system.

## II. INTEGRATED MANUFACTURING SYSTEM

In schematic diagram of the integrated manufacturing system in Fig. 1, input subsystem, management and control subsystem, manufacturing process subsystem, support and information processing subsystem, output subsystem and process design subsystem are shown along-with their symbols  $S_1, S_2, S_3, S_4, S_5, S_6$  respectively. The interconnections between subsystems  $i$  and  $j$  are represented as  $e_{ij}$  (e.g.,  $e_{21}$  represents connectivity/interaction/interdependence from subsystem  $S_2$  to  $S_1$ ).

Graph theoretic model for the integrated manufactur

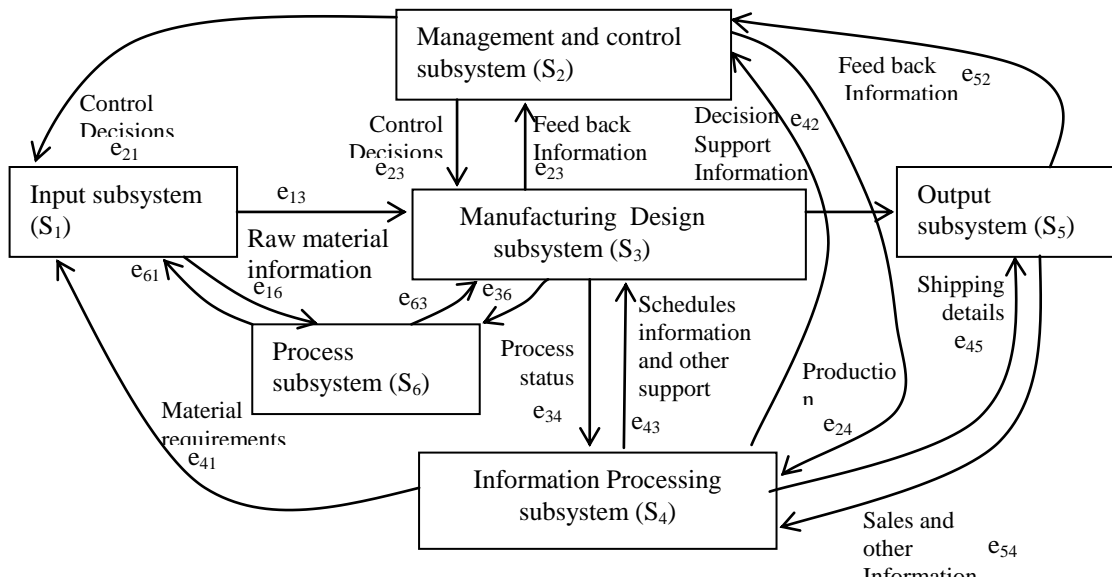


Fig. 1 Schematic diagram of integrated manufacturing system (Six subsystems)

ing system uses permanent matrix for the mathematical representation of individual subsystems as well as interactions among them. The permanent matrix of the integrated manufacturing system in Fig. 1 is developed as in equation (1) below.

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} S_1 \\ e_{21} \\ 0 \\ e_{41} \\ 0 \\ e_{61} \end{matrix} & \begin{bmatrix} 0 & e_{13} & 0 & 0 & 0 & e_{16} \\ S_2 & e_{23} & e_{24} & 0 & 0 & 0 \\ e_{32} & S_3 & e_{34} & e_{35} & e_{36} & 0 \\ e_{42} & e_{43} & S_4 & e_{45} & 0 & 0 \\ e_{52} & 0 & e_{54} & S_5 & 0 & 0 \\ 0 & e_{63} & 0 & 0 & 0 & S_6 \end{bmatrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} \end{matrix} \quad \text{---(1)}$$

The permanent function of the matrix P in equation (1) is developed as given below. The terms in the permanent multinomial of the integrated manufacturing system are following a specific pattern and are arranged in six different groups as shown in equation (2).

It can be observed from equation (2) that there are total 56 terms in the permanent multinomial of the integrated manufacturing system. In equation (2), the terms represent interaction procedures among the subsystems (e.g., the term  $S_1 S_4 e_{52} e_{23} e_{35}$ , represent interactions among three subsystems  $S_5, S_2$  and  $S_3$  while  $S_1$  and  $S_4$  represent standalone subsystems). The terms of the permanent multinomial in equation (2) above correspond to the real subsets of the manufacturing system structure and identify all possible structural patterns contributing towards various goals of the manufacturing system and offer different structural tests that may be used to analyze the integrated manufacturing system in different distinct ways.

$$\begin{aligned}
 \text{Per (P)} = & [S_1 S_2 S_3 S_4 S_5 S_6] + \\
 & [S_1 S_2 S_2 S_2 S_6 e_{54} e_{45} + S_1 S_2 S_2 S_3 S_4 e_{43} e_{34} + S_1 S_3 S_5 S_6 e_{24} e_{42} + S_1 S_4 S_5 S_6 e_{32} e_{23} + \\
 & S_2 S_2 S_2 S_4 S_1 e_{16} e_{61} + S_2 S_1 S_4 S_5 e_{63} e_{36}] + \\
 & [S_1 S_2 S_6 e_{54} e_{43} e_{35} + S_1 S_3 S_6 e_{24} e_{43} e_{32} + S_1 S_5 S_6 e_{42} e_{23} e_{34} + S_1 S_4 S_6 e_{52} e_{23} e_{35} + \\
 & S_4 S_5 S_6 e_{32} e_{21} e_{13} + S_1 S_3 S_6 e_{24} e_{45} e_{52} + S_2 S_5 S_6 e_{41} e_{13} e_{34} + S_2 S_4 S_5 e_{61} e_{13} e_{36}] + \\
 & \left[ \left\{ S_1 S_6 e_{42} e_{23} e_{35} e_{54} + S_2 S_3 e_{24} e_{42} e_{16} e_{61} + S_2 S_5 e_{54} e_{45} e_{16} e_{61} + S_4 S_5 e_{32} e_{23} e_{16} e_{61} \right\} + \right. \\
 & \left. \left\{ S_2 S_5 e_{43} e_{34} e_{16} e_{61} + S_1 S_5 e_{24} e_{42} e_{36} e_{63} + S_2 S_1 e_{63} e_{36} e_{54} e_{45} \right\} \right] + \\
 & \left[ \left\{ S_1 S_6 e_{42} e_{23} e_{35} e_{54} + S_1 S_6 e_{52} e_{23} e_{34} e_{45} + S_1 S_6 e_{24} e_{43} e_{35} e_{52} + S_5 S_6 e_{42} e_{21} e_{13} e_{34} \right\} + \right. \\
 & \left. \left\{ S_4 S_6 e_{52} e_{21} e_{13} e_{35} + S_2 S_6 e_{54} e_{41} e_{13} e_{35} + S_5 S_6 e_{24} e_{41} e_{13} e_{32} + S_2 S_5 e_{63} e_{34} e_{41} e_{16} \right\} \right] + \\
 & \left[ \left\{ S_4 S_5 e_{63} e_{32} e_{21} e_{16} \right\} \right] + \\
 & \left[ \left\{ S_6 e_{54} e_{45} e_{32} e_{21} e_{13} + S_3 S_1 e_{16} e_{61} e_{24} e_{45} e_{52} + S_5 e_{16} e_{61} e_{43} e_{32} e_{24} + S_2 e_{16} e_{61} e_{43} e_{35} e_{54} \right\} + \right. \\
 & \left. \left\{ S_5 e_{16} e_{61} e_{42} e_{23} e_{34} + S_4 e_{16} e_{61} e_{52} e_{23} e_{35} + S_1 e_{63} e_{36} e_{52} e_{24} e_{45} + S_5 e_{24} e_{42} e_{61} e_{13} e_{36} \right\} \right] + \\
 & \left[ \left\{ S_2 e_{54} e_{45} e_{61} e_{13} e_{36} \right\} \right] + \\
 & \left[ \left\{ S_6 e_{52} e_{24} e_{41} e_{13} e_{35} + S_6 e_{54} e_{42} e_{21} e_{13} e_{35} + S_6 e_{52} e_{21} e_{13} e_{34} e_{45} + S_5 e_{63} e_{34} e_{42} e_{21} e_{16} \right\} \right. \\
 & \left. \left\{ S_5 e_{63} e_{32} e_{24} e_{41} e_{16} + S_4 e_{63} e_{35} e_{52} e_{21} e_{16} + S_2 e_{63} e_{35} e_{54} e_{41} e_{16} \right\} \right] + \\
 & \left[ \left\{ e_{32} e_{23} e_{54} e_{45} e_{16} e_{61} \right\} + \left\{ e_{61} e_{13} e_{36} e_{24} e_{45} e_{52} \right\} + \right. \\
 & \left. \left\{ e_{16} e_{61} e_{43} e_{35} e_{52} e_{24} + e_{16} e_{61} e_{42} e_{23} e_{35} e_{54} + e_{16} e_{61} e_{45} e_{52} e_{23} e_{34} + e_{54} e_{45} e_{63} e_{32} e_{21} e_{16} \right\} + \right. \\
 & \left. \left\{ e_{63} e_{35} e_{54} e_{42} e_{21} e_{16} + e_{63} e_{35} e_{52} e_{24} e_{41} e_{16} + e_{63} e_{34} e_{45} e_{52} e_{21} e_{16} \right\} \right]
 \end{aligned}$$

### III. RESULTS AND ANALYSIS

The graphical representation of first to seventh group structural patterns is shown in Fig. 2 and these groups consist of 1, 0, 6, 8, 16, 16 and 9 terms each respectively. The interactions among subsystems are presented in the form of interaction loops which are the graphical representation of subsystems how they interact with each other. The detailed discussion on interaction procedures among subsystems of group V terms as shown in Fig. 3 is presented below. Two of the subsystems in each term remain isolated under this group. The interaction cycles are present among four subsystems in two distinct forms. In subgroup (i), four subsystems are interacting in the form of two two-subsystem dyads while in subgroup (ii); the four subsystems interact in a four-subsystem loop. The observations of various sub-graphs under group V has been discussed below under two subgroups.



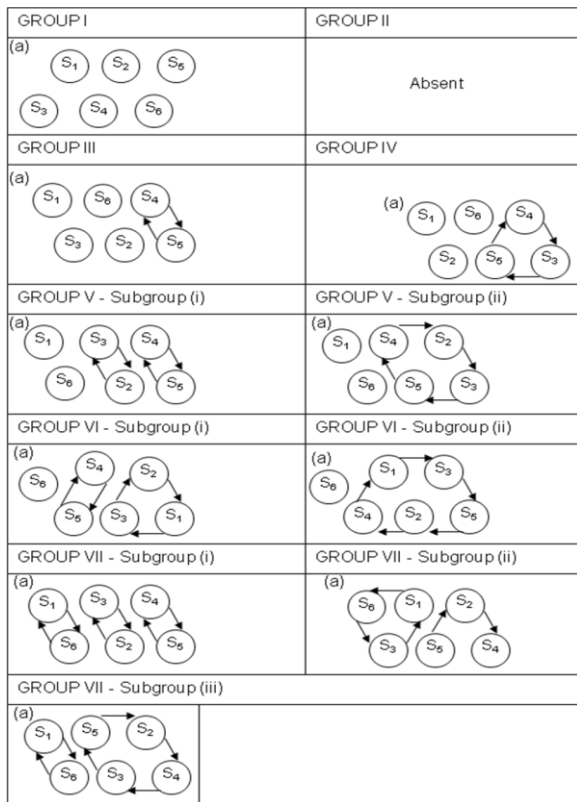


Fig. 2 Structural groups in integrated manufacturing system

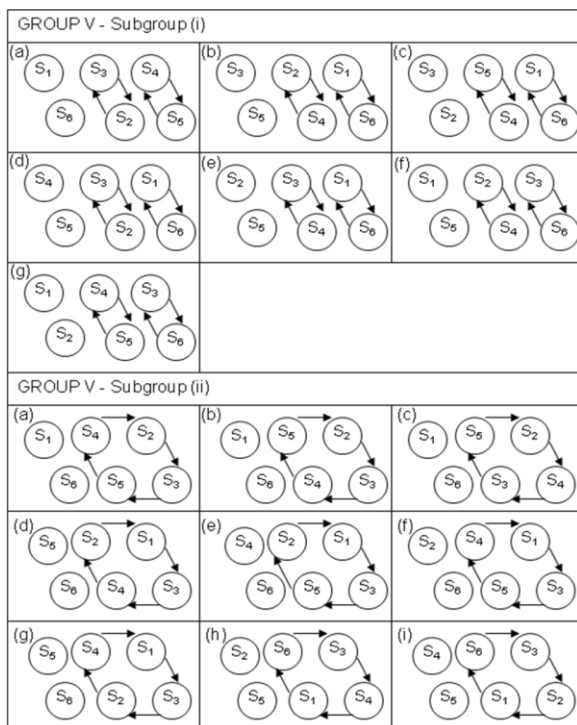


Fig. 3 Detailed members of group V in integrated manufacturing system

Sub group (i) of structural group V

- There are two interaction cycles between two subsystems each as well as a stand alone subsystem in subgroup (i) of group V. This mainly represents the simultaneous nature of two interaction cycles. In one of the interaction cycle, the link from S<sub>3</sub> to S<sub>2</sub> and vice-versa show the communication of strategic requirements. Meanwhile, other routine information exchange continues between S<sub>4</sub> and S<sub>5</sub>.
- In part (b), the interaction dyad between S<sub>2</sub> and S<sub>4</sub> represents the formal review process by S<sub>2</sub> of status of production system through reports generated from S<sub>4</sub> and then the updating action by S<sub>2</sub> on the schedule adjustments as deemed necessary. The interaction dyad between S<sub>1</sub> and S<sub>6</sub> show the simultaneous information exchange between input subsystem and the new design subsystem for finding the material availability or for vendor identification process.
- Two information exchange dyads in part (c) are between S<sub>4</sub> and S<sub>5</sub> as well as S<sub>1</sub> and S<sub>6</sub>. The later dyad represents the same activity as in part(b) while the former may be considered to be aimed at receiving routine information from the output subsystem about finished goods requirements as well as the communication of finished goods delivery schedule to the output subsystem.
- The part (d) also has two dyads, one between S<sub>2</sub> and S<sub>3</sub> which is representative of the direct feedback information received by S<sub>2</sub> from S<sub>3</sub> and then S<sub>2</sub> giving its advice to S<sub>3</sub> for improvements while the other between S<sub>1</sub> and S<sub>6</sub> representing the information exchange aimed at developing new product and new input requirements.
- The part (e) has two dyads of information exchange between S<sub>3</sub> and S<sub>4</sub> as well as S<sub>1</sub> and S<sub>6</sub>. Interaction dyad between S<sub>3</sub> and S<sub>4</sub> is aimed at updating the information subsystem with the manufacturing process status and also taking the scheduling and other required information from information subsystem. Simultaneously, S<sub>1</sub> and S<sub>6</sub> exchange information for new product developments and for the input requirements.
- The part (f) has two dyads of information exchange between S<sub>2</sub> and S<sub>4</sub> as well as S<sub>3</sub> and S<sub>6</sub>. Interaction dyad between S<sub>2</sub> and S<sub>4</sub> is aimed at receiving detailed updates by management subsystem from the information subsystem giving its own inputs also to information subsystem. Simultaneously, S<sub>3</sub> and S<sub>6</sub> exchange information for new product developments and for the manufacturing requirements for taking manufacturing concerns into consideration at design stage itself.
- The part (g) has two dyads of information exchange between S<sub>5</sub> and S<sub>4</sub> as well as S<sub>3</sub> and S<sub>6</sub>. Interaction dyad between S<sub>5</sub> and S<sub>4</sub> is aimed at giving detailed updates by information subsystem to the output subsystem on new products and receiving its

feedback. Simultaneously,  $S_3$  and  $S_6$  exchange information for new product developments and for the manufacturing requirements for taking manufacturing concerns into consideration at design stage itself.

#### Subgroup (ii) of structural group V

- The interaction cycle in part (a) of subgroup (ii) of group V corresponds to the updating of  $S_4$  by  $S_5$  and then  $S_2$  directs  $S_3$  for meeting the requirements. Finally  $S_3$  follows these directions and supplies finished products or delivery schedule information to  $S_5$ .
- The interaction cycle in part (b) of subgroup (ii) of group V represents the  $S_5$  sending special and urgent requirements as per customer priorities to  $S_2$  and  $S_2$  giving prioritization on orders to work directly to the  $S_3$ . Then  $S_3$  updates  $S_4$  of the real time status and finally  $S_5$  receiving updated information on action taken in real time on its requirements made earlier.
- The interaction cycle in part (c) of subgroup (ii) of group V is also of the similar nature as the previous one. Here also,  $S_5$  discusses the special requirements with  $S_2$  but in a non-urgent way. Then  $S_2$  makes required changes in the schedules and updates in  $S_4$  in a routine way. Finally,  $S_3$  utilizing updated plans to manufacture as per current needs.
- In the interaction cycle in part (d) of subgroup (ii) of group V, link from  $S_4$  to  $S_2$  shows the use of transaction data in  $S_4$  by  $S_2$  for taking various actions regarding  $S_1$  i.e., suppliers. The further link to  $S_3$  shows the assessment of effect of changes in  $S_1$  on  $S_3$  and finally  $S_3$  recording its observations on  $S_1$  into  $S_4$ .
- In this interaction cycle in part (e) of subgroup (ii) of group V, link from  $S_5$  to  $S_2$  shows the direct feedback from  $S_5$  to  $S_2$  pointing to some need for adjustment in  $S_1$ .  $S_2$  then gives its directions to  $S_1$  and then  $S_3$  utilizing the corrected inputs from  $S_1$  and then supplying the new outputs to  $S_5$ .
- The interaction cycle in part (f) of subgroup (ii) of group V shows routine material requirements schedule retrieval by  $S_1$  from  $S_4$  (e.g., MRP generated by  $S_4$  for dependent items) and thus keeping ready the raw material required to  $S_3$  proactively.  $S_3$  then supplying to  $S_5$  and again  $S_5$  updating  $S_4$  about the new emerging requirements in a recurring way.
- In the interaction cycle in part (g) of subgroup (ii) under group V,  $S_2$  takes direct feedback from  $S_3$  about various  $S_1$  components (e.g., vendors etc.).  $S_2$  then updates  $S_4$  database with vendor ratings and thus  $S_4$  then is able to select appropriate vendor for  $S_1$ .
- In part (h), the interaction loop is from  $S_6$  to  $S_3$  to  $S_4$  to  $S_1$  and finally back to  $S_6$ . This loop of information exchange may be viewed to be aimed at new design subsystem giving its initial input about new product designs to manufacturing process subsystem and

manufacturing process subsystem communicating it to input subsystem for possible vendor identification through information subsystem. New design subsystem receives the feedback from inputs about vendor queries and the cycle of information continuing unless all issues of product development are resolved.

- In part (i), the interaction loop is from  $S_6$  to  $S_3$  to  $S_2$  to  $S_1$  and finally back to  $S_6$ . This loop of information exchange may be viewed to be aimed addressing new product development issues and new design subsystem giving its initial input about new product designs to manufacturing process subsystem and manufacturing process subsystem apprising the management subsystem of its strengths/weaknesses with respect to the new design. Management subsystem then communicates with the input subsystem for possible vendor identification through information subsystem. process design subsystem receives the feedback from inputs about vendor queries and the cycle of information continuing unless all issues of product development are resolved.

#### IV. CONCLUSION

The present study has used comprehensive model to describe the interaction procedures among various subsystems of integrated manufacturing system through detailed observations and analysis for its continuous improvement. The following can be concluded from present observation and analysis.

- a) The different nature of activities of various subsystems during their interactions has also been discussed.
- b) The utility of various subsystems has been reported based on the analysis of graphical interactions among subsystems.
- c) Study also critically analyzes the pattern of information-exchange among various subsystems in various groups in the presence of isolated/standalone subsystems.
- d) This study may help in assessing the capability of the manufacturing system to address various functional needs. The objectives/ goals of specific integrated manufacturing systems can be achieved by identifying, analyzing and modifying various structural elements through the present investigation.

#### REFERENCES

- [1] Deo, N., 2000, Graph Theory, Prentice-Hall of India, New Delhi.
- [2] Jurkat, W.B. and Ryser, H.J., 1966, "Matrix Factorization of Determinants and Permanents", Journal of Algebra, Vol 3, pp01–27.
- [3] Mason-Jones, R., Berry, D. and Naim, M.M., 1998, "A systems engineering approach to manufacturing systems analysis" Integrated Manufacturing Systems, Vol. 9 No. 6, pp.350-365.
- [4] Minc, H., 1966, "Upper bounds for permanents of (0, 1) - matrices", Journal of Combinatorial Theory, Vol 2, pp 321-326.

- [5] Rao, R.V. and Gandhi, O.P., 2001, "Digraph and matrix method for selection, identification and comparison of metal cutting fluids", Proc. IME, Journal of Engineering Tribology, Vol. 212, pp. 307–318.
- [6] Rao, R.V. and Gandhi, O.P., 2002, "Digraph and matrix methods for machinability evaluation of work materials", International Journal of Machine Tools and Manufacture, Vol. 42, pp.321–330.
- [7] Singh, V. and Agrawal, V.P., 2008, "Structural modeling and integrative analysis of manufacturing systems using graph theoretic approach", Journal of Manufacturing Technology management, Vol 19 No. 7, pp.844-870.
- [8] Venkatasamy, R. and Agrawal, V.P., 1996, "Selection of automobile vehicle by evaluation through graph theoretic methodology", International Journal of Vehicle Design, Vol. 17 No.4, pp. 449-470.
- [9] Venkatasamy, R. and Agrawal, V.P., 1997, "A digraph approach to quality evaluation of an automobile vehicle", Quality Engineering, Vol. 9 No.3, pp. 405-417.
- [9] Singh, V., Agarwal V. and Deb, P., An Improved Graph Theoretic Model for Integrated Manufacturing System', National Conference on Recent Advances in Manufacturing (RAM-2010), Surat, India July 2010.

# JUST IN TIME MANUFACTURING AND INVENTORY CONTROL MANAGEMENT: A RETROSPECTIVE AND LITERATURE REVIEW

A. Puran singh B. Ritesh Kumar

## Abstract

This paper reviews a brief historical development for WIP Inventory Management in various systems. The aim is to study only the inventory control policies in JIT. The control policy affects performance of a manufacturing system and can be classified as push, pull, or combination of pull and push. A pull or just-in-time (JIT) production system is a philosophy or an approach of the manufacturing system in which order release occurs due to physical removal of finished inventory in response to the customer demand. In this paper, a review of behavior of a manufacturing system in terms of performance parameters under the control of different JIT techniques is discussed. The considered control policies are kanban, CONWIP, and hybrid which are based on planned elimination of all waste and continuous improvement of productivity. A separate comparison among all the control policies in terms of performance parameters has also been included in this study. At the end, a table summarizes the use of JIT strategy or its techniques such as kanban, CONWIP and kanban-CONWIP (referred as hybrid) in manufacturing systems from the internationally reputed researches.

**Keywords** – WIP Inventory control, JIT, MRP, ERP, kanban system, CONWIP, Hybrid system

## 1. A BRIEF HISTORY

The basic elements of JIT were developed by Toyota in the 1950's, Ohno (1988), and became known as the Toyota Production System (TPS). JIT was well-established in many Japanese factories by the early 1970's. JIT began to be adopted in the U.S. in the 1980's (General Electric was an early adopter), and the JIT concepts are now widely accepted and used. JIT is a concept for producing a required volume of required item at a required point of time. It is to meet out the global competition, in which the work-in-process inventory (WIP) is managed and controlled more accurately than the Material Requirement Planning (MRP)-production system to reduce the production cost (Golhar and Stamm 1991 and Monden 1983).

Prior to the dominance of the computer in manufacturing, inventory was controlled using reorder-point/reorder-quantity (ROP/ROQ) type methods.

A.Department of Mechanical &Automation Engineering, ASET Bijwasan  
Delhi, India

Puran.singh910@gmail.com

B.Department of Mechanical &Automation Engineering, ASET Bijwasan  
Delhi, India

During the 1960's, Joseph Orlicky, Oliver Wight along with others developed a new system, which they termed Material Requirements Planning (MRP). Orlicky et al (1975) obviously believed that they were on to something big; he subtitled his book on the subject "The New Way of Life in Production and Inventory Management"

(Orlicky 1975). After a slow start, MRP began to gather steam during the 1970's fueled by the "MRP Crusade" of the American Production and Inventory Control Society (APICS). Orlicky (1975) reported 150 implementations in 1971. By 1981, the number had grown to around 8,000 (Wight 1981). As it grew in popularity, MRP also grew in scope, and evolved in the 1980's into Manufacturing Resources Planning (MRP II), which combined MRP with Master Scheduling, Rough-Cut Capacity Planning, Capacity Requirements Planning, Input/Output Control and other modules. In 1984 alone, 16 companies sold \$400 million in MRP II software (Zais 1986). By 1989, over \$1.2 billion was sold to American industry, constituting just under one-third of the entire software industry (IE 1991).

While MRP was steadily dominating the American production control scene, the computer was far less pervasive in production and inventory control. Instead, several Japanese companies, most notably Toyota, developed the older ROP/ROQ methods to a high level. Starting in the 1940's, Taiichi Ohno began evolving a system that would enable Toyota to compete with American automaker but would not depend on efficiencies resulting from long production runs that Toyota did not have the volumes to support. This approach, now known as the "Toyota Production System," was designed to "make goods, as much as possible, in a continuous flow" Ohno (1988).

By the end of the 1980's, JIT began being eclipsed by the next great thing Enterprise Resources Planning (ERP). With the development of the client/server information technology architecture, it became feasible to integrate virtually all of a corporation's business applications with a common data base. ERP offered both near-total integration and "best of breed" software in the specific applications. Of course, ERP was much more complex than MRP II, containing modules for every business function imaginable, from accounting and financial functions to human resources. And it was correspondingly more expensive, with implementation costs at some companies soaring as high as \$250 million (Boudette et al 1999).

After that business/manufacturing philosophy goes by different names like: Agile manufacturing, lean manufacturing, synchronous manufacturing, world-class manufacturing, zero inventory production system, stock-less production system, and continuous manufacturing system (Cheng(1996) and Womack J.P, et al 1990).ERP also one continued to grow in popularity. As the 1990's drew to a close and fear of the Millennium Bug intensified, ERP was being installed at a feverish rate. As ERP began its rise, it appeared that the JIT movement had run its course. Even though Toyota continued to rank at the top of the automotive industry in quality and efficiency metrics, interest in the Toyota Production System was on the wane. But in 1990, a landmark case study conducted by MIT was published in *The Machine That Changed the World* by Womack and Jones and Roos (1990). This study compared American, European, and Japanese automobile manufacturing techniques and concluded in no uncertain terms that the Japanese methods, particularly those of Toyota, were vastly superior.

In addition, the authors freshened JIT by recasting it as "Lean Manufacturing." With a new name and a new set of stories to rekindle interest, the system created by Taiichi Ohno again became a hot topic in the world of manufacturing. The JIT movement also spawned a separate movement that ultimately became larger than JIT itself—Total Quality Management (TQM). Originally cast as a means for facilitating smooth production flow, TQM grew into a popular management doctrine institutionalized in the ISO 9000 Certification process. The focus on TQM in the 1980's also spurred Motorola to establish an ambitious quality goal and to develop a set of statistical techniques for measuring and achieving it. This approach became known as "Six Sigma" and was eventually adopted by companies such as Allied Signal and then General Electric. Six Sigma entered the mainstream management lexicon when Jack Welch, the charismatic CEO of GE, declared that it played a major role in his company's financial success. Today, Six Sigma carries on the legacy of TQM just like lean carries on the legacy of JIT. Despite traveling separate paths since emerging from JIT, it now appears that the Lean Manufacturing and Six Sigma movements are about to merge, as suggested by the recent best selling management book, *Lean Six Sigma* George, M.L et al (2002). At the same time, the computer approach characterized by MRP/ERP is also undergoing consolidation. After Y2K proved to be a non-event, ERP became passes as a term but was quickly replaced by SCM (Supply Chain Management). Remarkably, SAP, the world's largest provider of ERP software, changed its entire product line in a matter of months (or so it would appear when a search of their web-site revealed that all references to ERP were gone and had been replaced with the new SCM acronym). Since both ERP and SCM trace their roots back to quantitative production and inventory control, it is perhaps not surprising that they would wind up embodied in the same software products. However, all this leads one to wonder when the Lean Six Sigma Supply Chain Management movement will begin.

## 2. JIT AND INVENTORY CONTROL

While corporate America tried to navigate this sea of buzz words and to comprehend the Toyota Production System along with the mysterious concept of pull. As with any new management trend, academic researchers were quick to get on the JIT band wagon. It appears the first academic paper describing kanban was published by Japanese researchers in 1977. The title of the paper by Sugimori, et al.(1977) is telling: "Toyota Production System and Kanban System: Materialization of Just-In-Time and Respect-For-Human System." JIT manufacturing was closely associated with the principles of pull production control. This pull system is controlled by downstream information and is inherently make-to-stock. For example closed lines are pull systems because buffer spaces act as stock voids to trigger releases (Berkley 1992 & Gaury et. al. 2001). The objectives of pull system can be listed as:

- Producing the right part in the right place at the right time.
- Eliminating waste due to any activity that increased cost without adding value, i.e. unnecessary movements of materials, excess inventory, faulty production methods, and rework etc.
- Improve profits and ROI (Return on Investment) by reducing inventory levels increasing the inventory turnover rate, reducing variability, and improving product quality.
- To reach the goals of driving all inventory buffers toward zero by eliminating errors leading to defective items since there are no buffers of excess parts.
- Implement quality program, for supplier quality assurance, for stoppages. Workers, to understand the personal responsibility, to stop production when something goes wrong, to indicate line slowdowns or stoppages, and to record and analyze causes of production
- Stabilize and level the MPS (Master Production Schedule) with uniform plant loading by creating a uniform load on all work centers through constant daily production.
- Meet demand fluctuations through end item inventory rather than through fluctuations in production level times also allow economical production of smaller lots
- Try for single setup times or "one touch" setup through, better planning, process redesigning, and product redesigning, using specialized equipment. Single setup.
- Reduce lead times by moving work stations close together; applying group technology and cellular manufacturing concepts, reducing queue length, reducing delivery lead times through close cooperation with suppliers, and achieving the idle lot size of one unit.
- Use machine and worker idle time to maintain equipment and prevent breakdowns.
- To train workers to operate several machines, to perform maintenance tasks, and to perform quality inspections.
- Implementing the Toyota Production System concept of "respect for people" for a good relationship between workers and management.
- Use a control system such as kanban (card) system to convey parts between work stations in small quantities.

"JIT" focuses on abolishing or reducing Muda ("Muda", the Japanese word for waste) and on maximizing or fully utilizing activities that add value from the customer's perspective. From the customer's perspective, value is equivalent to anything that the customer is willing to pay for in a product or the service that follows. So the elimination of waste was the basic principle of JIT production system.

All the waste sources described in different literature are related to each other and getting rid of one source of waste can lead to either elimination of, or reduction in others. Perhaps the most significant source of waste is inventory, work in process and finished parts inventory do not add value to a product and they should be eliminated or reduced. When inventory is reduced, hidden problems can appear and action can be taken immediately. There are many ways to reduce the amount of inventory, one of which is reducing production lot sizes, reducing lot sizes however, should be followed by a set up time reduction so as to make the cost per unit constant as the famous economic order quantity formula states Abdullah Fawaz et al (2003). A reduction in WIP will reduce the number of defects in the event of a problem. JIT is not an inventory control system; it is a philosophy for continuous improvement of quality that puts emphasis on prevention rather than correction. Reduction of Inventory will also open up more space in the factory (24, Matt Schmidt).

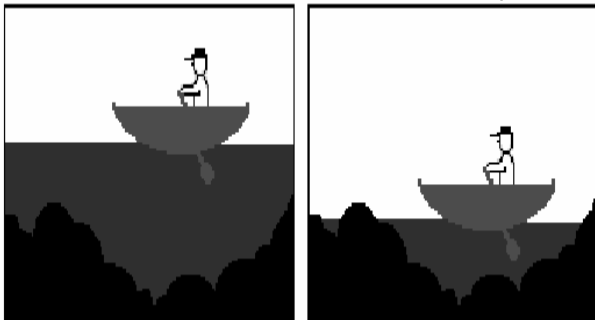


Figure 2.1

Next is an analogy of how reducing inventory will aid in overall performance. Inventory hides problems in a process, if reduce it, then problems in the system becomes clearer. The figure 2.1, in the first picture the boat is clearly above all rocks in the water, and no problems clearly exist. In the second picture the water level lower and the rocks have now become a problem for the boat Matt Schmidt et al. The water level represents inventory, the rocks represent problems in the system, and the boat represents the company's operations. If inventory is reduced, problems in the system become clearer from the operations standpoint [24].

The general idea of inventory in JIT is to establish flow processes by linking work centers so that there is even and balanced flow of materials throughout the entire production process (Al-Tahat and Mukattash 2006). Unfortunately pull systems do not lend themselves to all business types because of, product types, lead times and any stock holding arrangements with customers. However, there are so many

benefits by adapting JIT techniques, which are listed in Table I.

Table I: Pull system benefits

Reduces costs	Improves quality	Improves customer service	Maintains flexibility
•reduces average WIP	•improved defect detection	•short cycle times	•avoid direct congestion
•reduced space	•improved communication	•reduce sources of process variability	•less reliance on forecasts
•little rework		•remotes shorter lead times	•promotes floating capacity

With the above discussion we have come to these remarks that the traditional manufacturing methods have a target throughput which has to be specified and the actual throughput of the system has to be monitored which is not quite suitable for a production system in present scenario, however, controlling the amount of work-in-process or the finished goods inventory is more easy than that controlling the throughput or cycle time.

But the inventory-based control systems react to the changes in inventory level directly. This may leads over reacting to natural variation of the demand process instead of reacting only to the shifts in demand arrival rate. Therefore, a demand detecting mechanism is needed to determine whether a real change in demand rate occurs (Veatch and Wein 1994). Finally, it is noticed that traditional systems are also bad during execution than the JIT systems. Therefore, to meet customer expectations with on-time delivery of correct quantities of desired specification without excessive lead times or large inventory levels, pull production control is required. The pull control systems may also be further divided as kaban, CONWIP, Hybrid etc. on the basis of the sequence of order release, customer order arrival, material withdrawal and production, when to switch control, and where control is required. Thus, the following subsections describe the exhaustive reviews on Kanban, CONWIP, and Hybrid.

### 3. KANBAN MECHANICS

The classic version of kanban as pioneered by Toyota is called a "two card kanban" system, as shown schematically in Figure 3.1. Production begins when a material handler with a move card removes a standard container of parts from the outbound stock point. The move card authorizes that handler to take these parts and tells him where the parts are needed. Before the container is removed, the production card is removed from the standard container and placed on the production card board.2 Production cannot begin without a production card, a container of the appropriate incoming parts, and an idle work station. When all three are available, the worker removes a standard container of parts from the inbound stock point, removes the move card from the container, places it in the hopper for move cards at the work station and begins to process the parts. Periodically, a material handler collects the move card(s), locates the needed parts, transports them to the

workstation, and the process repeats at the next workstation upstream. It is easy to see that the two card kanban system is the result of an artificial distinction between parts processing and material movement. If we include material movement as a separate process, we see the two card kanban system becomes a one card system with the move card becoming a “production” card for the move process.

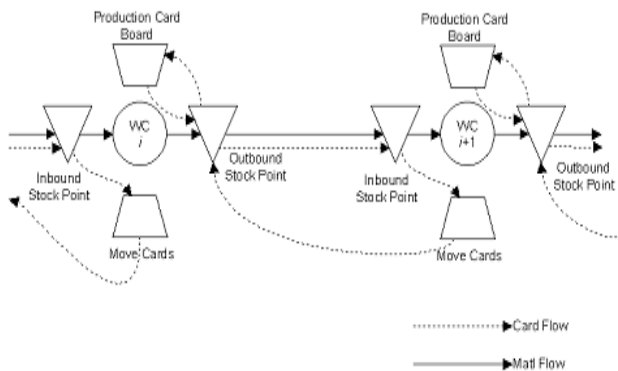


Figure 3.1

Dasci and Karakul (2008) presented a model to analyze a manufacturing system which is operating under pull-type control and shows pull production control is often implemented using kanban systems. Kanban control policy links production activities and transmitted demand information from finished buffers to the preceding workstation using cards called “kanban”. There are many implementation forms of Kanban, (Qi Hao and Shen 2008), model complex kanban based material handling system in an assembly line using both discrete event and agent-based technologies through hybrid simulation approach. Perros and Altiok (1986) described a Kanban controlled unreliable manufacturing system in which the machine failure and repair rates were assumed to follow exponential distributions. Material flow in the system was controlled by manufacturing blocking discipline. Kanban system especially in the upstream stages, may not respond quickly enough to changes in the demand.

#### 4. CONWIP SYSTEM

Another considered inventory control policy in this research is CONWIP which is a generalized form of kanban and initially proposed as a pull alternative to kanban (Spearman et. al. 1990). It is such a policy where a raw part enters to the system after servicing of a finished part to the customer in response of a demand. The aim of CONWIP is to combine the low inventory levels of Kanban with the high throughput of MRP System. CONWIP also shared the benefits of kanban such as shorter lead times and reduced inventory levels while being applicable to a wide variety of production environments.

#### 5. HYBRID (KANBAN- CONWIP)

Much research has been done on individual control systems, only few comprehensive hybrid studies exist i.e. Generalized Kanban control proposed by Buzacott and Hanifin (1978) based on kanban and base stock control

policies. In CONWIP policy, inventory levels are not controlled at the individual stages hence high inventory levels building up in front of bottleneck stages. Bonvik et. al. (1997), proposed hybrid policy which is a combination of Kanban-CONWIP to reduce loose coordination between production stages in a CONWIP lines. Hybrid policy can be implemented as a straightforward modification to a kanban policy, simply by routing kanbans from the finished goods buffer to the first production stage instead of the last.

#### 6. COMPARISON OF THE CONTROL POLICES IN JIT

Several researches demonstrate comparisons between kanban & CONWIP considering various performance parameters of manufacturing line. Reviews on pull systems also showed that few comparison studies have compared performance of CONWIP and hybrid (kanban-CONWIP) and kanban CONWIP and hybrid systems through simulation, experimental, analytical models and case studies. With the conclusion of the theoretical statements and simulation study of CONWIP, Spearman et. al. (1990), proposed that the CONWIP system can be used by any manufacturing system where the utility of kanban system is limited. This shows the superiority of CONWIP pull system is an alternative to kanban system.

Yang (2000) compared different kanban and CONWIP system and showed that kanban produces the longest mean customer waiting time with high WIP. Gaury et.al. (2000), described a methodology using evolutionary algorithm and discrete-event simulation for the choice of a pull production-control strategy and model Kanban, CONWIP, and Hybrid lines with six, eight, and ten stages. In a flow line model based on an actual system in a Toyota assembly factory, Bonvik et. al. (1997) showed the comparison in some specific situations. While comparing the production policies, the hybrid control policy demonstrated superior performance in achieving a high service level target with minimal inventories, closely followed by CONWIP. The performance measures used are: (i) service level or fill rate (ii) amount of inventory or WIP. Deterministic demand situation is assumed. Cases were considered including both constant and time-varying demand rates. Spearman and Zazanis (1992), showed that CONWIP produces a higher mean throughput than Kanban. In the same scenario, Muckstadt and Tayur (1995) showed that CONWIP produces a less variable throughput and a lower maximal inventory than Kanban. In a survey paper, Framinan et. al. (2003), discussed operations and applications of different CONWIP production control systems with detailed comparisons. Takahashi et. al. (2005) applied Kanban, CONWIP and synchronizes CONWIP to supply chains in order to determine the performances of a system. They considered supply chains containing assembly stages with different lead times. Geraghty and Heavey (2005) also presented a comparison of the performance of several pull-type production control strategies in addressing the service level v/s WIP trade-off in an environment with low variability and a light-to-medium demand load. Gstettner and Kuhn (1996), found that Kanban achieved a given throughput level with less WIP than CONWIP. Hodgson and Wang (1991) presented strategy where the first two

stages 'push' and all other stages 'pull'. They did not compare the different control policies and showed only the results of this hybrid combination.

Paternina and Das (2001) applied a optimization technique called Reinforcement Learning (RL) and a heuristic policy named Behavior-Based Control (BBC) on a four-station serial line. The numerical results were used for comparison of control policies such as CONWIP, kanban and other hybrid policies on the basis of total average WIP and average cost of WIP with two different (constant and Poisson) demand arrival processes. Duri et. al. (2000) and Geraghty & Heavey (2004) compared policies in a different scenario for a specific automobile assembly line.

## 7. COMPARISON IN TERM OF PERFORMANCE PARAMETERS

For comparing of different inventory control policies in terms of the performance of a manufacturing system, various performance parameters have been considered in several research papers, Gupta and Gupta (1989), concluded that high production rates can be realized only when the number of Kanbans is chosen optimally. Framinan et. al. (2006) have been established the correct number of cards in pull systems that can be addressed either statically (i.e. card setting), or dynamically (i.e. card controlling). They reviewed the different contributions regarding card controlling in pull systems (especially for CONWIP) and then a new procedure was proposed and tested under different environments. Philipoom et. al. (1987), described factors that influence the number of kanbans required in implementing JIT production techniques. They include throughput, process variation, machine utilization, and processing times. Takahashi (2003) and Takashashi & Nakamura (2002), proposed a reactive control mechanism for Kanban control system. The system adjusted the amount of Kanban cards according to a detected change in demand process using the time series data of the finished goods inventory level. Chan (2001) presented the effect of kanban size on various parameters i.e. in process inventory, service level or fill rate, unsatisfied order, manufacturing lead-time. Kern et. al. (1996) examined the relative effectiveness of various rescheduling policies by a simulation experiment in JIT manufacturing environment. They analyzed schedule instability, total units of sales lost and average finished inventory. Alabas et. al. (2002) found that the Tabu search requires less computational efforts when compared to genetic algorithm (GA), simulated annealing (SA) and the neural network meta-model. They have used algorithms to find the optimum number of kanbans with the minimum cost by a simulation model. Tang et. al. (1993) used Taguchi method in the simulation experiments to study the relationship between the multiple performance measures and some given dispatching rules. The parameters which they used are utilization, number of machines buffer size and work in process for the operations planning and scheduling problems in FMS. Tardif and Maaseidvaag (2001) introduced a new adaptive kanban-type pull control mechanism which determined the timings to release or reorder raw parts based on customer

demands and inventory back orders, in order to maximize marginal benefits for predicting steady state performance of a manufacturing system. Petroni and Rizzi (2002) used average WIP, average Flow (production lead) time, Mean Tardiness as performance measures for predicting performance of a manufacturing system. Shahabudeen et. al. (2002 & 2003) designed single and two cards dynamic kanban systems using a simulated annealing algorithm. They proposed a universal test based on percentage zero demand (PZD), mean lead-time (MLT) and mean total WIP (MTW) and may be suited for the MOP (measure of performance) in any JIT system. Koukoumialos and Liberopoulos (2005) have been developed a general purpose analytical approximation method for the performance evaluation of a multi-stage, serial, echelon kanban control system. An iterative procedure was used to determine the unknown parameters of each subsystem. Jing (2003) presented about the improving the performance of job shop manufacturing by reducing setup/processing time variability. He has measured three factors for shop performance which are average work-in-process (WIP) inventory, average flow time and average set up time to processing time ratio.

## REFERENCES

- [1] Abdullah Fawaz, Lean Manufacturing tools and Techniques in the process industry with a focus on steel, university of Pittsburgh, 2003.
- [2] Alabas, C.; Altiparmak, F.; Dengiz B. (2002). A comparison of the performance of artificial intelligence techniques for optimizing the number of kanbans, *Journal of Oper. Res. Society* 53 (8), 907-914
- [3] Takahashi, K. (2003). Comparing reactive kanban systems, *Int. J. Prod. Res.*, 41(18), 4317-4337
- [4] Takahashi, K.; Myreshka; Hirotoni, D. (2005). Comparing CONWIP, synchronized CONWIP, and kanban in complex supply chains, *Int. J. Prod. Econ.*, 93-94, 25-40
- [5] Takahashi, K.; Nakamura, N. (2002). Decentralized reactive kanban system, *Eur. J. Oper. Res.*, 139(2), 262-276
- [6] Tang, L. L.; Yih, Y.; Liu, C. Y. (1993). A study on decision rules of a scheduling model in an FMS, *Computers in Industry*, 22(1), 1-13
- [7] Tardif, V.; Maaseidvaag, L. (2001). An adaptive approach to controlling kanban systems, *Eur J. Oper. Res.*, 132(2), 411-424
- [8] Tersine, RichardJ; Principles of inventory and materials management, Elsevier
- [9] Veatch, M. H.; Wein, L. M. (1994). Optimal control of a two-station tandem production/inventory system, *Operations Research* 42(2), 337-350
- [10] Womack, J.P., D.T. Jones and D. Roos, *The Machine That Changed the World: The Story of Lean Production*, HarperCollins Publishers, New York, 1990.
- [11] Wight, O., "Input/Output Control a Real Handle on Lead Time," *Prod. & Inv. Mgmt. J.* 11:3,9-31, 1970. Just-in-Time: Proceedings of the APICS Zero Inventories Seminar, Orlando, FL: APICS (1983). science publishing co, Inc., 1988.
- [12] Womack, J.P, et al, *The machine that changed the world* (Macmillan publishing company, Canada, 1990).
- [13] Wight, O., *MRP II: Unlocking America's Productivity Potential*. Boston: CBI Publishing, 1981.
- [14] Yang, K. K. (2000). Managing a flow line with single-Kanban, dual- Kanban or CONWIP, *Production and Oper. Management*, 9(4), 349-366.





# FFT BASED EVALUATION OF CUTTING FORCES AND CHATTER VIBRATIONS IN TURNING BY VARYING SPEED, FEED, DEPTH OF CUT AND RAKEANGLE.

A. N. B. Gandhi .

**Abstract--** The prediction of chatter vibrations between the tool and work piece is important as guidance to the machine tools user for an optimal selection of depth of cut and spindle rotation, resulting in maximum chip removal rate without the undesirable vibration. This can be done by some approaches. In this work, an experimental method is applied in which the time-varying directional dynamic turning forces coefficients are expanded in Fourier series and integrated in the width of cut bound by rake angles. The forces in the contact zone between tool and work piece during the cut are evaluated by an algorithm using FFT and a dynamometer located between the work piece material and the cutter geometry. The modal parameters of the machine – work piece – tool system like natural frequencies, velocity, acceleration and roller bearing must also be identified experimentally. At this point, it is possible to plot the frequency, velocity, acceleration to this dynamic system. these curves relates the spindle speed with axial depth of cut, separating stable and unstable areas, allowing the selection of cutting parameters resulting maximum productivity, with acceptable surface roughness and absence of chatter vibrations. Experimental FFT turning tests were performed in the machine, using a four single point cutting tool with different rake angle. The result showed perfect agreement between chatter prediction and experimental tests. We have performed 200 experiments on four tools using FFT analyzer and HMT lathe machine. FFT has generated 10,000 row of Data. The detailed analysis of FFT spectrum and data of it is used for conclusion.

**Keywords--** turning, cutting forces, chatter vibrations, FFT Analyzer ,DFT signals ,spectrum

## I. INTRODUCTION

Vibration concerns the repetitive motion of an object or objects relative to a stationary frame referred to as the equilibrium of the vibration. Vibrations may be measured in terms of displacement, velocity or acceleration.

Vibrations exist everywhere and may have a great impact on the surrounding environment. One general phenomenon of vibration is the "self-oscillation" or resonance, meaning that a system exposed to even a weak force, which excites a resonance, may result in a substantial vibration level that eventually results in damage to or failure of the system. Thus, it is of great importance in engineering design to

consider the properties of the system from a vibration point of view, referred to as the dynamic properties of the system. Note: Simple Harmonic Motion is a periodic motion. It happens when the restoring force is proportional to displacement.

The turning operation is a cutting process using a single point cutting tool. producing periodic chip and an impact when the edge touches the work piece. The case is heated and stressed during the cutting part of the cycle, followed by a period when it is unstressed and allowed to cool. The consequences are thermal and mechanical fatigue of the material and vibrations which may be explained by two distinct mechanisms, called "mode coupling" and "regeneration waviness", explained in Tobias (1965), Koenig Berger & Tlusty(1967) and Budak & Altintas (1995). The mode coupling chatter occurs when forced vibrations are present in two directions in the plane of cut. The regenerative chatter is a self excitation mechanism associated with the phase shift between vibrations waves left on both sides of the chip and happens earlier than the mode coupling chatter in most machining case, as explained by Altintas(2000). In Turning one of the machine tool works piece system structural modes is initially excited by cutting forces. The wavy surface left by a previous cutting is removed during the succeeding, which also leaves a wavy surface due to structural vibrations. The cutting forces become oscillatory whose magnitude depends on the instantaneous chip dynamic thickness, the inner and outer chip surface. The cutting force can grow until the system becomes unstable and the chatter vibrations increase to a point when the cutter jumps out of the cut or cracks due to the excessive force involved. These vibrations produce poor surface finishing, noise and reduce the life of the tool. In order to avoid this undesirable effect the speed, feed rate, depth of cut and rake angle are chosen at conservative values, reducing the productivity.

There are many parameters which impact vibration such as cutting speed, feed rate, tool geometry and defect, and machine rigidity. Much research has been done in this field, with most based on simulations or mathematical modeling. This paper developed an experimental approach to analyze the generated by turning with measurement data processed

by fast Fourier transform (FFT). The FFT is mainly used to transform time domain signals into the frequency domain. However, in this paper, the FFT is applied to the space domain signal.

Now a days, the FFT is used in many areas, from the identification of characteristic mechanical vibration frequencies to image enhancement standard routines are available to perform the FFT by computer in programming languages such as PASCAL, FORTRAN & C and many spread sheet and other software packages (like Mat lab) for the analysis of numerical data allow the FFT of a set of data values to be determined readily.

#### A. Experimental setup

The vibration of shafts manufactured by turning was measured in the test by dynamometer and FFT analyzer. VIB EXPERT hardware was used for collecting all vibration signature and numerical data of signature and spectrum of experiment.. OMNITREND SOFTWARE was used for export all VIBEXPERT data to computer. The measurements were parallel to the feed direction (the shaft axis) with the magnetic sensors, one on head stock spindle & other on tail stock spindle nearer to work pieces side.

The vibration of the work piece was measured on work pieces manufactured by turning using various feed rates, Speed rates, depth of cut and rack angle. The cutting tools (H.S.S) was used with shaft material (M.S.) turned on lathe with dynamometer. Hence the machine had different vibration characteristics during turning. The turning was conducted with feed rates: 0.05 mm/r, 0.1 mm/r, 0.21 mm/r, and 0.42 mm/r speeds 124 rpm, 194 rpm, 392 rpm & 512 rpm & depth of cut 0.5 mm, 1 mm, 1.5 mm with four different tools with rakeangle 14°, 8°, 4° & 0° (r is the abbreviation for revolution)



## II. DATA PROCESSING

Excel sheet was used to analyze the profiles and numerical data. The measured data (.mdb file) was directly loaded by the program as input, with the results outputted mostly in the form of signature, spectrum and numerical data from FFT.. All the loaded data was treated at the beginning to make the analysis easier and more reliable. The data was processed first in a normalization step-sorted and then in a shift step-filtered. In the first step, the data was sorted as minimum vibration and maximum MRR. Then in the second step the data was filtered as different rake angle-14°, 8°, 4° & 0°

The FFT is an efficient algorithm for computing the discrete Fourier transform (DFT) of a sequence. The FFT returns the DFT of signals, computed with a fast Fourier transform algorithm. In the algorithm, the FFT decomposes the long series of Fourier transform into a short series of Fourier transforms, using the fastest Fourier transform in the West (FFTW) Mat lab.

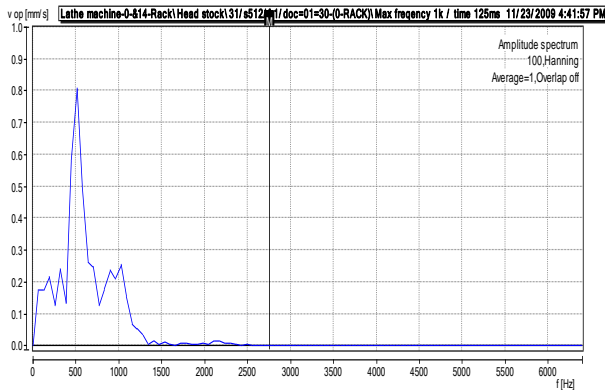
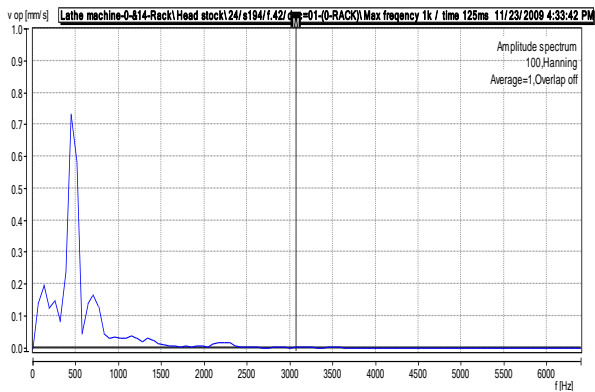
## III Results and discussion

The work piece vibration generated by rakeangle turning is affected by the feed rate, macro tools geometry, speed, depth of cut, and machine tools vibrations, etc. The frequency for any cutting condition, which is the reciprocal of the feed rate. The high

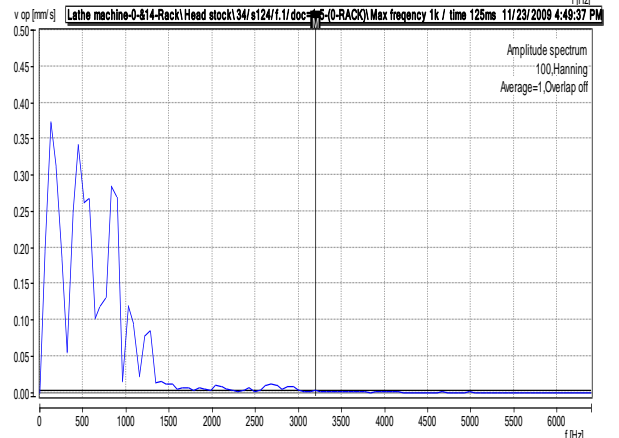
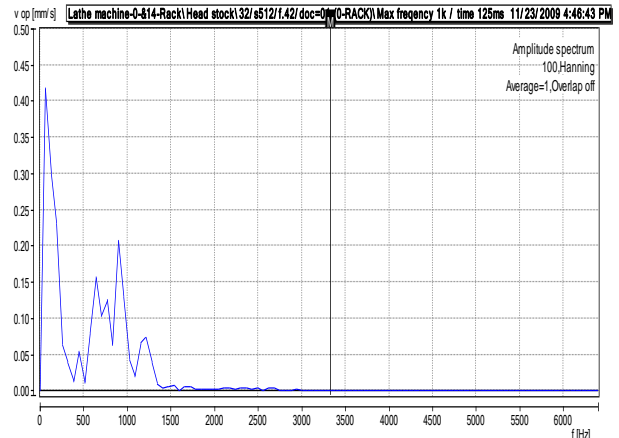
frequency can be related to the micro tool geometry (tool wear) and the machine tool vibrations. Since the experiments were conducted with a new cutting tool, the differences caused by the only rack angle considered other cutting tool geometry will be ignored in the discussion.

The frequency plots of the surface profiles generated by the 2 rack angles 2 speed of at are shown in figure 1-4 for various feed rates.

**Fig. 1** frequency feature of the profile for a feed rate of 0.21 mm/r , speed 512 rpm, & depth of cut = 1.0mm tool rack angle = 4 and 14 rack angle =8



**Fig. 2** frequency feature of the profile for a feed rate of 0.21 mm/r , speed 512, & depth of cut = 1 mm tool rackangle = 4 and rack angle = 8



**Fig. 3** frequency feature of the profile for a feed rate of 0.21 mm/r , speed 124, & depth of cut = 1mm tool rack angle = 4 and rack angle = 8

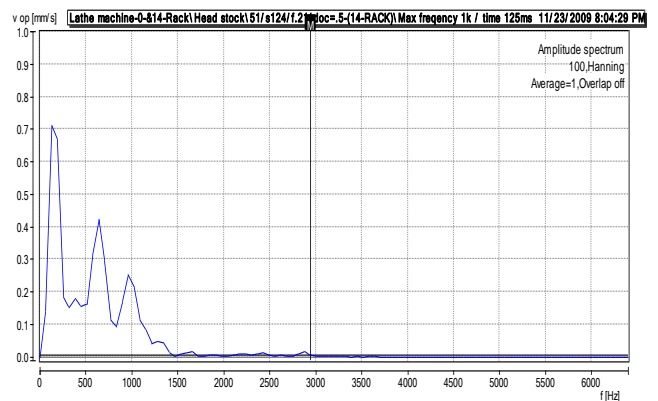
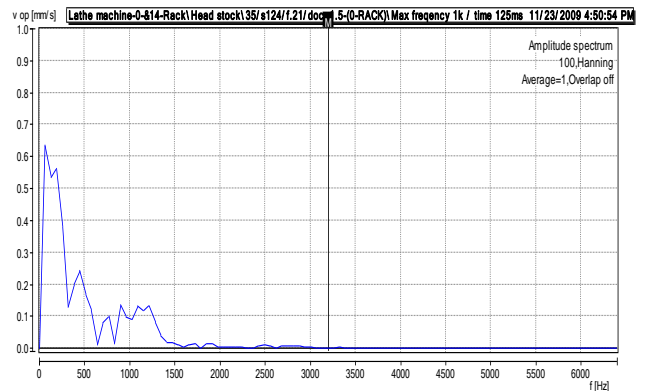
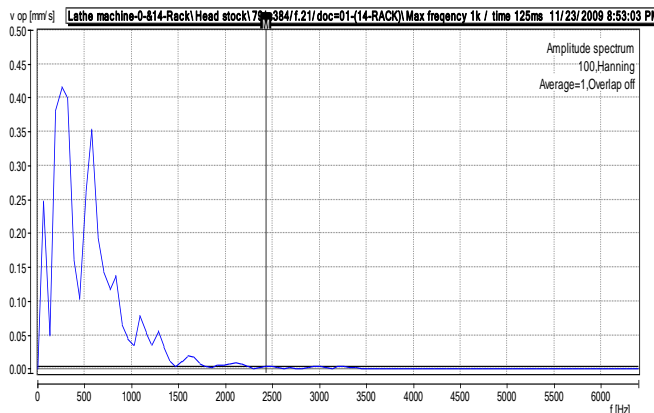
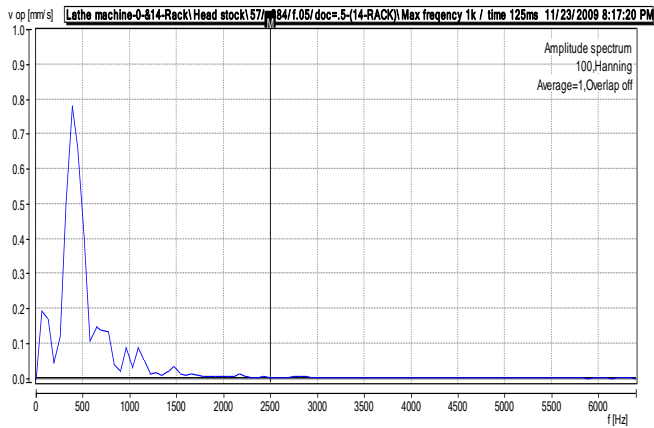


Fig. 4 frequency feature of the profile for a feed rate of 0.21 mm/r , speed 394, & depth of cut = 1.0mm tool rack angle = 4 and rack angle = 8



The plots shows that the vibration generated by 0° rack angle tool have less high frequency components than the ones generated by the 14° rack angle tool especially at high feed rates (0.42 mm/r), which reflects machine.

At the low feed rate (0.05 mm/r), the frequency features of the vibration were on the two quite different rack angle of machine tool 14° rack angle tool the feed frequency (1/0.05 = 20) does not have a large amplitude while many high frequency components appear at irregular frequencies.

4. CONCLUSION

The main advantage of the vibration measurement through the FFT is the metal removal rate maximization, at the same time avoiding the adverse effect of chatter, the poor surface finish, noise and breakage of tools. The FFT analysis the extent of the effect of rack angle of tools on the machine vibrations at various feed rates. At low feed, rates (0.05 mm/r), 124rpm speed 0° rakeangle tools significantly influenced the machine and decreases the stability of

vibration. At medium or high feed rates (> 0.21mm/r), the machine vibration had little impact on the vibration with the effect of vibration at high feed rates (> 0.42 mm/r) being considerable vibration compared to other factors such as tool wear. Note that these frequency features and the effect of machine vibrations at these feed rates & speed are strongly related to the machine conditions. For example, if two other Rack angle were used, the vibrations may have more impact at the feed rate of 0.1 mm/r.

The effects of tool rake angles on the stability of self-excited primary chatter vibration in high-speed oblique turning process are experimentally investigated. The experimental results indicate that a larger rake angle increases the stability of vibration and reduces the feed force. The results also reveal that the excitation energy is provided by the friction force on the flank surface and the cutting force acting on the tool rake. And when feed rate and depth of cut also increase with rack angle decreases the stability of vibration

TABLE :1 EFFECT OF DIFFERENT FEED RATES ON FEED FORCES

( speed 124RPM )(0.5mmDOC)

FeedRate (mm/rev)	Feed forces (Kg)			
	0°Rack	4°Rack	8°rack	14°Rack
0.05	4.16	2.08	2.08	06.25
0.10	6.25	4.186	4.18	10.41
0.21	8.33	8.85	8.33	06.25
0.42	12.33	17.5	11.41	10.41

TABLE :2 EFFECT OF DIFFERENT SPEED ON FEED FORCES

(Feed 0.21mm/rev) ( 1mm DOC)

Speed (rpm)	Feed forces (Kg)			
	0°Rack	4°Rack	8°rack	14°Rack
124	8.33	14.37	9.58	7.18
194	28.75	11.97	7.18	7.18
384	16.66	11.97	9.58	11.97
512	10.41	9.58	7.18	9.58

REFERENCES

1.LI Haosheng, Wu SU , HuBERT KRATZ- FFT and wavelet based analysis of the influence of machine vibration on hard turned surface topology- vol.12,no:4,Aug:2007-International Journal of machine tool and manufacture

2. Abele E, fiedler U Creating stability lobe diagrams during turning Annals of CRIP 53(1)-309-312
3. FFT Spectrum Analyser applet guidance notes.
4. H.B.Lacendra and V.T.Lima -Evaluation of Cutting Forces and Prediction of Chatter Vibration in milling- vol.26,No.1, March-2005
5. [Hamdan, M. N.](#); [Bayoumi, A. E.](#), An approach to study the effects of tool geometry on the primary chatter vibration in orthogonal cutting, Journal of Sound and Vibration, Volume 128, Issue 3, p. 451-469.2003

# Effect of Fuel properties on Performance of Downdraft Gasifier

A. Hardik B Kothadia, B. Rajesh N Patel

**Abstract--** It observed that very little information on parametric influence of gas generation rate, fuel variation and particles size variation on performance and temperature level of downdraft gasifier is available in literature. While studying a performance of gasifier, the choice of fuel is important for efficient and economical use of gasifier, based on this realization This research work presents the result of the experimental investigations carried out using different biomass to investigate the fuel properties on gasifier performance in downdraft gasifier. For this purpose three different fuel as coal, charcoal and wood are selected and performance of gasifier are measured base on gas generation rate calorific value of producer gas and gasifier efficiency in downdraft gasifier. It is observed that gasifier performance is highly dependent on fuel properties.

**Index terms -** Gasification; Fuel properties; Downdraft gasifier.

## I. INTRODUCTION

Gasification is a century old technology, which flourished quite well before and during the second world war. The technology disappeared soon after the Second World War, when liquid fuel became easily available. The interests in the gasification technology have undergone many ups and downs in running century [1]. Today, because of increased fuel prices and environmental concern, there is renewed interest in this century old technology. Gasification has become more modern and quite sophisticated technology.

The advantage of this technology is decentralized energy conversion system which operates economically even for small scale. A gas producer is a simple device consisting of usually cylindrical container with space for fuel, air inlet, gas exit and grate. It can be made of fire bricks, steel or concrete and oil barrels. The design of gasifier depends upon type of fuel used and whether gasifier is portable or stationary. Gasifier alone itself is of little use. The complete gasification system consists of gasification unit (gasifier), purification unit and energy converter - burners or internal combustion engine.

The key to a successful design of gasifier is to understand the properties and thermal behavior of the fuel as fed to the gasifier. Operation of gasification system demands knowledgeable and skilled operator. Those interested in this

technology must remember that it requires hard work and tolerance. Compared to conventional system such as liquid fuel run engines, gasification technology is inconvenient. But it is economical at many places and may lead to self-reliance in the crucial time of fuel crisis.

## II. Gasification Process

The gasifier is a reactor that converts any organic material into clean gaseous fuel called producer gas. The gasifier is essentially a chemical reactor where various complex physical and chemical processes take place. Four distinct processes take place in a gasifier as the fuel makes its way to gasification. They are: (1) Drying of fuel (2) Pyrolysis- a process in which tar and other volatiles are driven off (3) Combustion (4) Reduction. Distribution of heat in each stage of process, movement of composition of reaction in each stage and compositions of producer gas is shown in Fig.1

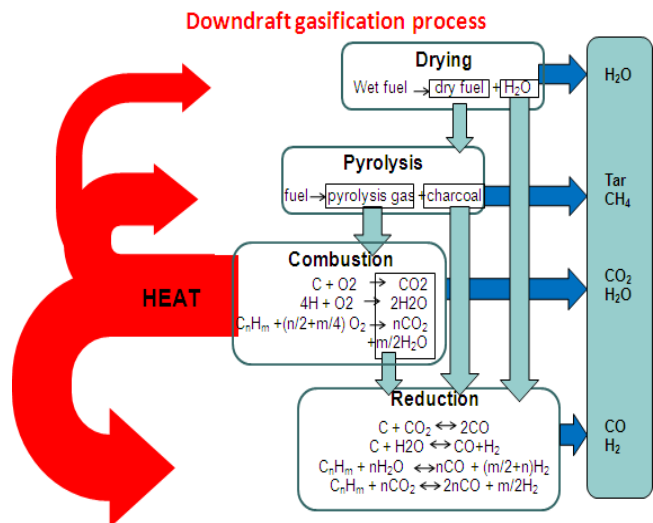


Fig.1 Downdraft gasification process

### A. Drying

All moisture must be removed from the wood chips in order to successfully produce wood gas. While all water in the wood will be vaporized eventually by heat of pyrolysis, combustion, and reduction, failure to remove moisture from the wood beforehand results in the inability to produce clean fuel. Therefore, in an ideal gasifier, some of the heat produced during combustion is used to completely dry the wood. Since the lack of clean gas could be catastrophic to our work, we have chosen to either pre-bake or purchase pre-dried wood chips.

A. Assistant Professor, Mechanical Engg. Department, Gandhinagar Institute of Technology, Gandhinagar e-mail: [h\\_kothadia@yahoo.in](mailto:h_kothadia@yahoo.in) )

B. Head of Department, Mechanical Engg. Department, Institute of Technology, Nirma University, Ahmedabad (e-mail: [rnp@nirmauni.ac.in](mailto:rnp@nirmauni.ac.in) )

### III. Materials and Methods

#### B. Pyrolysis

Pyrolysis occurs when the wood chips are heated without enough oxygen to burn. Fast pyrolysis does not begin until the wood has reached a minimum temperature of about 550 K [2]. This causes the wood chips to decompose into tars (gasses and liquids) and charcoal. The tars are burned off, leaving charcoal with high carbon content. This charcoal is integral for the reduction process.

#### C. Combustion

Combustion is the exothermic combination of hydrocarbons with oxygen. The heat for all processes is generated from combustion of the tars produced during pyrolysis. In addition, combustion produces carbon dioxide and hydrogen gas, which will become reactants in the reduction step. Adequate mixing and high temperature flame is key, since the lack of either could permit the tars to still be present in the wood gas, which in turn could lead to engine failure. Therefore, producing clean fuel is largely dependent on the combustion dynamics in the gasifier. The United Nations Forestry and Agriculture Organization suggest that the combustion zone should reach a minimum of 1473 K to ensure a clean burning fuel [2].

#### D. Reduction

Reduction reverts completely combusted hydrocarbons into a form that can be used as fuel. Note that reduction is the opposite of combustion - it is the endothermic removal of oxygen from hydrocarbons. Typically combustion and reduction exist in equilibrium in any burning process. Reduction in a gasifier occurs when carbon dioxide and water vapor flows through heated charcoal (primarily carbon). The heated carbon removes the oxygen from both the carbon dioxide and the water vapor. The oxygen is spread to the carbon atoms, forming covalent bonds in the form of carbon monoxide. Oxygen has a higher affinity for carbon than either hydrogen or itself. This leaves the remaining hydrogen atoms to form their natural diatomic. Therefore, two reactions occur from the addition of carbon and heat, carbon dioxide is reduced to carbon monoxide and water vapor is reduced to diatomic hydrogen and carbon monoxide. The study found that the rate of reduction is only high enough to run a gasifier at temperatures above about 973 K [2].

#### E. Post-Reduction

After the reduction process, the products are ready for combustion once again. These gases will be drawn into the cylinders of an internal combustion engine. There, they will be mixed with air, compressed, and then ignited to produce power for the go-kart. The additional oxygen present in the air will react with carbon monoxide to form carbon dioxide and with hydrogen to form water vapor as products, which will be our primary exhaust components.

#### A. Experimental procedure

It always ensures that the gasifier system (gasifier, cyclone and surge tank) is clean before starting any new experiment. It also ensures that there is no choking within the system through the pressure drop across cyclone, orifice and surge tank. The prepared char is fed into the gasifier up to air tuyeres then fuel (biomass) is fed into the gasifier and the top cover is closed. All the air tuyeres are closed except the firing nozzle. Fuel is fired with the help of mashal from the firing nozzle when the inner part seems to be red hot mashal removed. Then firing nozzle is closed and air tuyeres are opened. Flow rate is maintained with the help of U-tube manometer at design level and temperature along the length of gasifier is recorded. Thereafter the flow rate and temperature is recorded in regular interval. Meanwhile material is fed up to have continuous operation. Gas is collected for analysis in gas balloon from one of pet cock.

The procedure is repeated by changing fuel or particles size. The position of ball valve and opening of air tuyeres are kept same throughout experiment and for all the experiment with different fuel and particles size.

#### B. Char preparation

Char is a material of fuel being used for reduction material. It is made from same material which used in feed stoke for experiment of gasifier on lignite it is advisable to prepared char from lignite. Take require quantity (as per experimental set up feed char up to throat) of biomass. Fire the biomass with the help of diesel and allow it to burn for 20minuts. When observe that it catches fire all over, cover it with suitable vessel and kept it for almost 2.5 hour. The char is blackish in color and it shine.

#### C. Experimental set-ups

The schematic diagram of downdraft gasifier system is as shown in Fig.2. The components of gasifier system are as below:

- Gasifier
- Cyclone
- Orifice meter
- Surge tank
- Ball valve
- Blower
- Burner

In order to measure the pressure drop across each element the pet cock are inserted at 50mm of distance on inlet and outlet respectively. The temperature in gasifier is measure with eight thermocouples at regular distance.



Orifice is installed at 10 times the diameter of pipe after cyclone. The volume of surge tank is kept 100 times the volume of pipe in order to avoid chock up. Ball valve is placed below and after blower which maintains the flow rate of the system. At end the burner is placed where the producer gas is burnt. The flow line of system is explained as under.

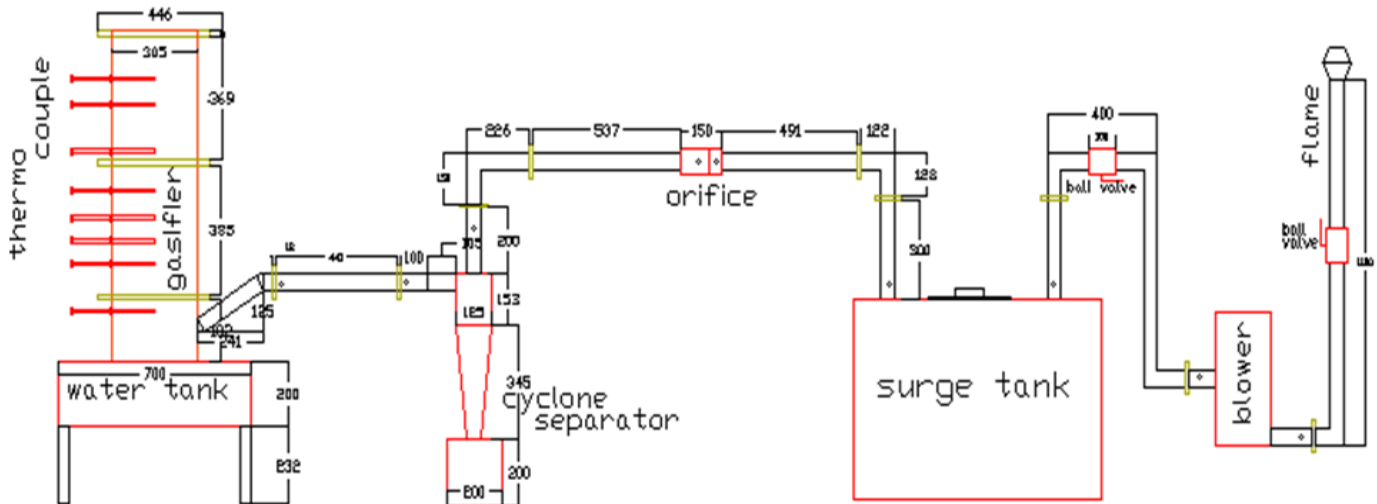


Fig.2 The schematic diagram of downdraft gasifier system

#### IV. RESULT AND DISCUSSION

Experiments were conducted on 10kWe downdraft gasifier with three different fuels coal, charcoal and wood to check the effect of fuel properties on gasifier performance.

A wide range of biomass fuels such as wood, charcoal, wood waste (branches, roots, bark, saw dust) as well agricultural residues- maize cobs, coconut shells, cereal straws, rice husks, can be used as fuel for biomass gasification. Theoretically, almost all kinds of biomass with moisture content of 5-30% can be gasified; however, not every biomass fuel leads to the successful gasification to check that our development work is carried out with common fuels such as coal, charcoal and wood. Key to a successful design of gasifier is to understand properties and thermal behavior of fuel as fed to the gasifier. The properties of fuel which influence the gasification are described below.

- Moisture content
- Particle size and distribution
- Bulk density of the fuel
- Volatile matter content
- Ash content and composition
- Energy content of fuel
- Fuel form
- Reactivity of fuel

Gas is generated in gasifier and is passing through the cyclone separator where the tar content is removed. Then the gas is passed through surge tank where is furthered cleaned and temperature is also reduce. Before the surge tank gas is passed through orifice plate. After surge tank it goes to the burner via blower where it burned.

##### A. Moisture content

The moisture content of the most biomass fuel depends on the type of fuel, it's origin and treatment before it is used for gasification. Moisture content of the fuel is usually referred to inherent moisture plus surface moisture. The moisture content below 15% by weight is desirable for trouble free and economical operation of the gasifier. Higher moisture contents reduce the thermal efficiency of gasifier and results in low gas heating values. Igniting the fuel with higher moisture content becomes increasingly difficult, and the gas quality and the yield are also poor.

##### B. Particle size and distribution

The fuel size affects the pressure drop across the gasifier and power that must be supplied to draw the air and gas through gasifier. Large pressure drops will lead to reduction of the gas load in downdraft gasifier, resulting in low temperature and tar production. Excessively large sizes of particles give rise to reduced reactivity of fuel, causing start-up problem and poor gas quality.

Acceptable fuel sizes depend to certain extent on the design of gasifier. In general, wood gasifier work well on wood blocks and wood chips ranging from 80x40x40 mm to 10x5x5 mm. For charcoal gasifier, charcoal with size ranging from 10x10x10 mm to 30x30x30 mm is quite suitable.

### C. Bulk density of fuel

Bulk density is defined as the weight per unit volume of loosely tipped fuel. Bulk density varies significantly with moisture content and particle size of fuel. Volume occupied by stored fuel depends on not only the bulk density of fuel, but also on the manner in which fuel is piled. It is also recognized that bulk density has considerable impact on gas quality, as it influences the fuel residence time in the fire box, fuel velocity and gas flow rate.

### D. Volatile matter content of fuel

Volatile matter and inherently bound water in the fuel are given up in pyrolysis zone at the temperatures of 100-150 °C forming a vapor consisting of water, tar, oils and gases. Fuel with high volatile matter content produces more tar, causing problems to internal combustion engine. Volatile matters in the fuel determine the design of gasifier for removal of tar. Compared to other biomass materials (Wood: 72-78 %, Coal: up to 40 %), charcoal contains least percentage of volatile matter (3-30 %)

### E. Ash content of fuel

A mineral content of fuel which remains in oxidized form after combustion of fuel is called ash. In practice, ash also contains some unburned fuel. Ash content and ash composition have impact on smooth running of gasifier. Melting and agglomeration of ashes in reactor causes slagging and clinker formation. If no measures are taken, slagging or clinker formation leads to excessive tar formation or complete blocking of reactor. In general, no slagging occurs with fuel having ash content below 5 %. Ash content varies fuel to fuel. Wood chips has contains 0.1% ash, while charcoal contains high amount of ash (16-23%)

### F. Energy content of fuel

Energy content of fuel is obtained in most cases in an adiabatic, constant volume bomb calorimeter. The values obtained are higher heating values which include the heat of condensation from water formed in the combustion of fuel. The heating values are also reported on moisture and ash basis. Fuel with higher energy content is always better for gasification. The most of the biomass fuels (wood, straw) has heating value in the range of 10-16 MJ/kg, whereas liquid fuel (diesel, gasoline) posses higher heating value.

### G. Reactivity of fuel

Reactivity determines the rate of reduction of carbon dioxide to carbon monoxide in the gasifier. Reactivity depends upon the type of fuel. It has found that wood and charcoal are more reactive than coal. There is relationship between reactivity and the number of active places on the char surfaces.

## V. CONCLUSION

Following are the major conclusions on the basis of above study.

- The fuel having low moisture content is desirable for trouble free and economical operation of the gasifier.
- Wood gasifier work well on wood blocks and wood chips ranging from 80x40x40 mm to 10x5x5 mm. For charcoal gasifier, charcoal with size ranging from 10x10x10 mm to 30x30x30 mm is quite suitable. In general, Acceptable fuel sizes depend to certain extent on the design of gasifier.
- The fuels having higher bulk density give higher quality gas and hence high efficiency of gasifier.
- Fuel with high volatile matter content produces more tar, causing problems to internal combustion engine.
- Fuels having low ash content and ash composition give smooth running of gasifier and also less causes of slagging and clinker formation.
- Fuel with higher energy content is always better for gasification.
- Fuels having high rate of reduction give higher calorific value gas. It has found that wood and charcoal are more reactive than coal.

## VI. REFERENCES

- [1] <http://zeep.com/zeep-technology/> accessed on 4th March 2012.
- [2] Avdhesh Kr. Sharma, "Experimental study on 75 kWth downdraft (biomass) gasifier system", *Renewable Energy* 34 (2009) 1726–1733.
- [3] Juan J. Hernández, Guadalupe Aranda-Almansa, Antonio Bula, "Gasification of biomass wastes in an entrained flow gasifier: Effect of the particle size and the residence time", *Fuel Processing Technology* 91 (2010) 681–692.
- [4] Francisco V. Tinaut, Andrés Melgar, Juan F. Pérez, Alfonso Horrillo, "Effect of biomass particle size and air superficial velocity on the gasification process in a downdraft fixed bed gasifier", *Fuel Processing Technology* 89 (2008) 1076–1089.
- [5] Zainal Z.A., Ali Rifau, G.A. Quadir, K.N. Seetharamu, "Experimental investigation of a downdraft biomass gasier", *Biomass and Bioenergy* 23 (2002) 283 – 289.
- [6] S.R. Patel, P.R. Bhoi, A.M. Sharma, "Field-testing of SPRERI's open core gasifier for thermal application", *Biomass and Bioenergy* 30 (2006) 580–583.
- [7] Catharina Erlich, Torsten H. Fransson, "Downdraft gasification of pellets made of wood, palm-oil residues respective bagasse: Experimental study", *Applied Energy* (2010).
- [8] T.B.Reed and A. Das' "Handbook of biomass downdraft gasifier engine system" Golden, Colorado, Solar Energy Research Institute, 1998.



# Optimization of SLG filler and flexural strength of composites cured in oven and microwave.

A. Jayant Vadher, B. Parita Sheth.

**Abstract**— A commercial phenol formaldehyde based resole thermosetting resin supplied by Borden Chemical Australia Pty. was filled with ceramic-based fillers (Envirospheres or SLG) to increase its flexural strength. By performing flexural tests at a range of filler addition levels, the optimal addition level of SLG was able to be determined in terms of workability, cost and performance. The composites obtained were post-cured in conventional oven and in microwaves respectively. It was found that the maximum flexural strength of the oven cured composites were only 5% higher than those cured in microwave when the percentage by weight of SLG was 24%. However, the time required for post-curing was also reduced from 10 hours (in conventional oven) to 40 minutes (in microwaves).

**Index Terms**— envirospheres, microwaves, Phenol formaldehyde, phenolic resin.

## I. INTRODUCTION

In the previous and concurrent studies, the flexural strength and flexural modulus of a commercial resole phenol formaldehyde resin reinforced with ceramic microsphere (SLG) filler, post-cured conventionally were determined using flexural tests. The filler percentage by weight is varied from 0 to 35%. Fillers will reduce shrinkage during molding, lower cost and improve strength. They are also used to improve electrical and thermal insulating properties and chemical resistance. This research is aimed at investigating the flexural strength of a commercial resole phenol formaldehyde resin reinforced with ceramic microsphere filler post-cured in microwaves conventional oven. The commercial resole resin used in this study was Hexion Cellobond J2027L. The acid catalyst used to crosslink the resin was Hexion Phencat 15. The ratio by weight of the resin to hardener for all samples in this work was chosen to be 50: 1. SLG (E-spheres) is a commercial ceramic microsphere product obtained as a fly ash by-product. The

particle size of this general purpose Espheres ranges from 20 – 300  $\mu\text{m}$  with approximate mean of 130  $\mu\text{m}$ .

The three point bending test provides values for the flexural modulus of elasticity in bending, flexural stress, flexural strain and the flexural stress-strain response of the material.

The main advantage of a three point flexural test is the ease of the specimen preparation and testing. However, this method has also some disadvantages: the results of the testing method are sensitive to specimen and loading geometry and strain rate. The standard used is ISO 14125:1998(E) because the results can then be compared with the work of others. A Material Testing Systems (MTS) 810 was used for the tests. The dimensions of the specimens of resins were 100mm x 10mm x 4mm and tested at a crosshead speed of 1 mm/min.

The material properties of greatest importance in microwave processing of a dielectric are the complex relative permittivity  $\epsilon = \epsilon' - j\epsilon''$  and the loss tangent,  $\tan \delta = \epsilon'' / \epsilon'$  the real part of permittivity,  $\epsilon'$ , sometimes called the dielectric constant, mostly determines how much of the incident energy is reflected at the air-sample interface, and how much enters the sample. The most important property in microwave processing is the loss tangent,  $\tan \delta$  or dielectric loss, which predicts the ability of the material to convert the incoming energy into heat. For optimum microwave energy coupling, a moderate value of  $\epsilon'$ , to enable adequate penetration, should be combined with high values of  $\epsilon''$  and  $\tan \delta$ , to convert microwave energy into thermal energy. Microwaves heat materials internally and the depth of penetration of the energy varies in different materials.

The tests has been performed to find properties like flexural strength, young's modulus and deflection which are quite necessary to observe the material properties.

II. SPECIMEN PREPARATION

The specimens were made 0 % to 35 % by weight of SLG in step of 5 % to cure phenol formaldehyde composite PF/E-SPHERES (X %), where X is the percentage by weight of the filler. As the raw materials of the composites are liquid and ceramic hollow spheres, the flexural test specimens were cast to shape. The resin is mixed with the catalyst, after which the SLG is added to the mixture and they are then mixed to give the uncured composite. The mould was made from PVC (poly vinyl chloride) sheets with six pieces of flexural test specimen each. After preliminary curing (minimum 72 hrs), the

%	Resin (g)	Hardener(g)	Filler (g)	Weight (g)
0	196.08	3.92	0	200
5	186.28	3.72	10	200
10	176.48	3.52	20	200
15	166.67	3.33	30	200
20	156.87	3.13	40	200
25	147.06	2.94	50	200
30	137.25	2.75	60	200
35	127.45	2.55	70	200

samples were taken out of the mould and Post-cure process has been done in conventional oven and microwave cavity after the removal of specimen from the mould. Oven temperature and time are as follows:

- 4 hours at 50 °C
- 4 hours at 80 °C
- 2 hours at 100 °C

Another set of specimen were post-cured in a modified SANYO microwave oven until the specimens reached approximately 100°C; this was achieved by setting up microwave time for 40 minutes with a power level of 80 W. Higher power levels were not recommended because it would cause the samples to cure swiftly, resulting in the formation of blowholes. Temperature of the specimens was measured using an infrared thermometer. 100°C was chosen because when samples were cured in a conventional oven, the highest temperature reached was 100°C. They were then cooled to

room temperature in the modified microwave oven cavity and then subjected to flexural tests.

III. TEST PROCEDURE

During the test, the upper support contacted the top surface of the test specimen, applying a load, which forced the piece to bend. This force was increased until the test specimen failed, at which point the maximum load (Peak Load (N)) that the specimen supported was recorded and the deflection at the mid point was noted. From this peak load, the flexural strength (or flexural stress) (see ISO14125) of the material can be found using the following equation (1) and was measured in mega Pascals (Mpa). From this, calculations of Flexural strain and then Young’s modulus of Elasticity in Bending (E) were conducted. Please note it is possible to use other values of load at different points of deflection to find the stress at a given point, though this research only dealt with load at failure.

Flexural Stress

$$\sigma_f = (3FL) / (2bh^2)$$

Flexural Strain

$$\epsilon_f = (6Dh) / (L^2)$$

Young’s Modulus,

$$E = (L^3m) / (4bh^3) = \sigma_f / \epsilon_f$$

The ISO 14125:1998 Fibre-reinforced plastic composites – Determination of flexural properties was used for the tests and some points to note are:

1. With the help of Vernier callipers, test specimens were measured and ensured that they assembled to required dimensional accuracy immediately after the specimen preparation.
2. The test system was then set up for the appropriate loading rate of 2mm per minute on the control-center computer attached.
3. Test specimen then placed on the supporting cylinders and the test was then commenced. At this time computer attached to the test system recorded real time data about load, stress, strain and modulus of elasticity.

4. After testing was completed, all data was gathered together and reviewed. In this review process, any test pieces that did not produce reasonable results were discarded.

IV. RESULTS AND DISCUSSION

1. Flexural strength

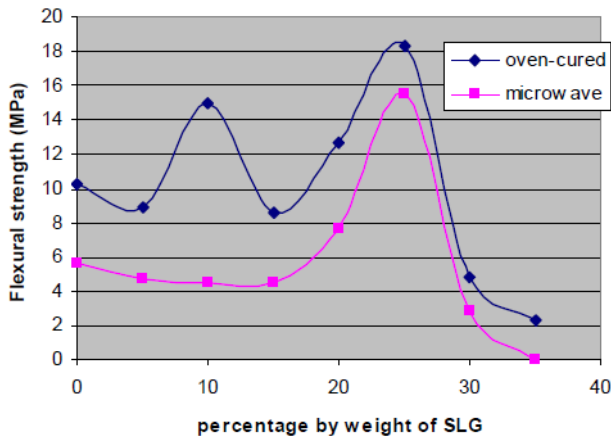


Fig. 1. Comparison of flexural strength of SLG reinforced phenolic composites.

It is clear from the figure that the flexural strength of, 0% by weight of SLG, conventionally cured specimen is much higher than that of the microwave-cured one; for the other percentages, the flexural strengths of conventionally cured specimens are approximately 64% to 92% higher than those of the microwave-cured ones. It is shown in the graph that highest flexural strength is 18.29 MPa for 25% at the oven cured condition which is 17.15% more than that of the microwave-cured one. For the 35% by weight of SLG, the flexural strength is not measurable for the specimens which were cured in microwaves because they were too low.

2. Flexural strain

The figure shows the comparison of strain induced in the specimen during testing and for conventional oven, the flexural strain of unfilled resin was highest and its value was 0.02515 mm/mm, which then dropped to its lowest of 0.0086 at 10% by weight of reinforcer. It rises again to a maximum of 0.0179 at 25% by weight of SLG before falling back sharply to 0.005125 at 35% by weight of filler.

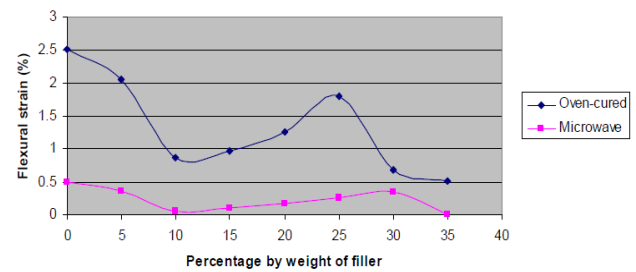


Fig. 2. Comparison of flexural strain of SLG reinforced phenolic composites.

For the microwave cured specimen the trend of the curve is in line with that of the flexural strength and it can be argued that the trend is correct. The flexural strain of unfilled resin was highest and its value was 0.01775 mm/mm, which then dropped to its lowest of 0.0025 at 15% by weight of reinforcer. It rises again to a maximum of 0.01783 at 30% by weight of SLG before falling back sharply to too low value at 35% by weight of filler the values of flexural strain mentioned above with their standard deviation. The flexural strain, at maximum flexural strength of 15.47 MPa, i.e. 25% by weight of SLG, was only 0.0116. So, at the highest flexural strength, flexural strain is comparatively more than that of the conventional cured specimen in microwave cured specimen.

3. Young’s modulus

Figure 3 shows the Young’s modulus of varying by weight of E-sphere (SLG) reinforced phenol formaldehyde matrix composite cured in oven. At zero percent weight of the SLG, the Young’s modulus is 409.54 MPa; it then increased to 430.35 MPa at 5 % by weight of SLG reinforcement; after this the values increase steadily to a high of 1736.05 MPa at 10% by weight of filler. The trend of the curve of Figure 3 differs with that of Figure 1 (flexural strength) and the maximum values of flexural strength and Young’s modulus does not occur at same percentage by weight of SLG.

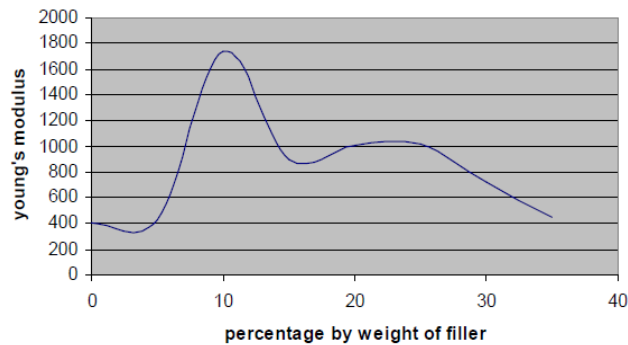


Fig. 3. Young's modulus of SLG reinforced phenolic composites (conventional oven).

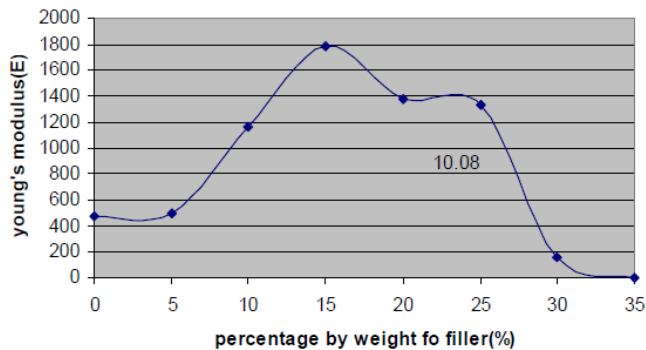


Fig. 4. Young's modulus of SLG reinforced phenolic composites (microwave).

Figure 4 shows the Young's modulus of varying by weight of E-sphere (SLG) reinforced phenol formaldehyde matrix composite cured in microwaves. At zero percent by weight of the SLG, the Young's modulus is 477.28 MPa; it then increased to a highest of 1163.64 MPa at 15 % by weight of SLG reinforcement; after this the values decrease steadily to a low of 157.08 MPa at 30% by weight of filler. Again, the trend of the curve of Figure 4 differs from that of Figure 1 (flexural strength) and the maximum values of flexural strength and Young's modulus does not occur at same percentage by weight of SLG.

## V. CONCLUSION

This study has evaluated the flexural strength, flexural strain and Young's modulus of varying percentage by weight of SLG reinforced phenolic resin composite post-cured in microwave irradiation; the best composite is the one with 25% by weight of SLG because the three properties above are either highest or quite high and the cost of the composite has been reduced by 24% but the flexural strength has been

increased by 180% however, the conventionally post-cured composites are preferred to those post-cured in microwaves because of higher mechanical properties and this is in opposition to the case of fracture toughness, in which the microwave post-cured composites are preferred (Ku et al., 2008c). It can be argued that when the fusion between phenolic resin (matrix) and SLG (reinforcer) is improved by adding some other fillers, e.g. garamite and resin to composite, its flexural strength will be further improved (Ku et al., 2008b).

It has been found that time required to cure the specimen in microwave is quite less than the specimen cured in conventional oven and it also supports greener environment but because of non-uniform heat transfer and localized hot spots specimen cured in microwave shows bit lower flexural strength in compare to that of specimen cured in conventional oven.

The result obtain by this research is little bit different from the research has been done by researcher means flexural strength comes quite smaller than the previous research which may be because of the specimen has been tested without grinding them from its sides to remove the porosity and so may result into the lower strength value.

## REFERENCES

- [1] Bahr, A, "microwave non-destructive testing methods", 1982.
- [2] Callister, W, *Fundamentals of Materials Science and Engineering*, John Wiley & Sons, New York, 2001, pp. 90
- [3] Clarke, J L (Editor), "*Structural design of polymer composites*", E & FN Spon, U.K., (1996), pp.59-62, 343-5, 357.
- [4] Ku, H S, "*Productivity Improvement of Composites Processing through the Use of Industrial Microwave Technologies, Progress In Electromagnetics Research*", PIER 66, (2006), pp.267-285..
- [5] ISO, *ISO 14125:1998(E), Fibre reinforced plastic composites – Determination of flexural properties, International Organization for Standardization, Geneva, Austria (1998)*.
- [6] Redjel, B, "*Mechanical Properties and Fracture Toughness of Phenolic Resin, Plastics, Rubber and Composites Processing and Applications*", 1995, Vol. 24, pp. 21-228.
- [7] Shackelford, J F, "*Introduction to Materials Science for Engineers*", 3<sup>rd</sup> edition, Macmillan, 1992, pp.435-437.
- [8] Smith, W F and Hashemir, J, "*Foundations of Material science and Engineering*", 4th edition, McGraw-Hill, (2006), pp. 523-525.
- [9] Strong, A B, "*Plastics: Materials and Processing*", 3rd edition, Pearson/Prentice-Hall, (2006), pp. 182-183, 304-309, 323-333, 620-621





# Effect of tube diameter on thermal performance of automobile radiator by the CFD analysis

A. P. K. Trivedi, B. N.B.Vasava

**Abstract:** The thermal performance of an automotive radiator plays an important role in the performance of an automobile's cooling system and all other associated systems. The flow behavior and temperature profile prediction in the radiator tubes are very useful information and is of great importance to the designer. The geometry of the tube type heat exchanger is an intricate one and there are no analytical optimization schemes available to optimize their design, while experimental trial and error is far too time consuming. The radiator designs at present depend on the empirical methods, wherein existing experimental data is used as the thumb rules for the design process. However, for any preliminary design the performance of the radiator can be accessed through Computational Fluid Dynamics (CFD) in prior to the fabrication and testing. In the current study a tube of an existing radiator is analyzed for evaluating the fluid flow and heat transfer characteristics. This paper presents a Computational Fluid Dynamics (CFD) modeling simulation of mass flow rate of fluid passing from the tubes of an automotive radiator. An introduction to mass flow rate and its significance was elaborated in order to understand the complications involved in the research and thereafter arrive at the objectives. Knowing the geometry of tube in radiator is the crucial application of CFD to numerically model and thereby analyze the simulation. The fluid flow simulation is conducted using commercial software ANSYS 12.1. The pressure and temperature distribution along the tube length and tube width are presented and analyzed. The results obtained serve as good database for the future investigations.

**Index terms:** Heat transfer, CFD, radiator design, Understood Simulation, Thermal management etc.

## I. INTRODUCTION

Increasing demands on engine power and performance to meet ever increasing customer requirements of safety against the burning car have necessitated improving the heat management system of the vehicle. The thrust on automobile manufacturers for developing compact and energy efficient cars warrants a thorough optimization process in the design of all engine components, including radiators. Radiators are installed in automobiles to remove heat from the under hood which include engine cooling and heat removal during air-conditioning process.

A. Student, M.E-II (Cad/Cam), Dept. of Mechanical Engg., Government Engineering College, Dahod, Gujarat-389151, (email-[pravintrivedi@hotmail.com](mailto:pravintrivedi@hotmail.com))

B. Student, M.E-II (Cad/Cam), Dept. of Mechanical Engg., Government Engineering College, Dahod, Gujarat-389151, (email-[vasavanaran@rediff.com](mailto:vasavanaran@rediff.com))

Manufacturers of commercial vehicles are facing a substantial increase of heat release into the cooling system. The main sources for this increase are: more stringent emissions leading to new combustion technologies and increased power of the engines. The total increase in cooling requirement may be up to 20% over the current level. At the same time the noise levels have to decrease and the fuel economy have to be improved. This forces the manufacturers to think about new concepts and optimized efficiency of the cooling system. CFD Analysis of the radiator as a whole is seldom reported in the literature. However, the technical information about fin-tube heat exchangers is available in bits and pieces as individual analysis carried out experimentally and/or numerically.

Hilde Van Der Vyer et al. [1] conducted a CFD simulation of a three dimensional tube-in-tube heat exchanger using Star-CD CFD software and made a validation test with the experimental work. The authors were fairly successful to simulate the heat transfer characteristics of the tube-in-tube heat exchanger. Though this work is not directly related to our study, this has been used as the base for the procedures of CFD code validation of a heat exchanger.

Witry et. al., [2] carried out CFD analysis of fluid flow and heat transfer in patterned roll bonded aluminum radiator, in which FLUENT's segregated implicit 3-D steady solver with incompressible heat transfer is used as the tool. In this study, shell side airflow pattern and tube side water flow pattern are studied. The authors presented the variation of overall heat transfer coefficients across the radiator ranging from 75 to 560 W/m<sup>2</sup>-K. This study established the capability of FLUENT code to handle such problems.

Chen et al., [3] made an experimental investigation of the heat transfer characteristics of a tube-and-fin radiator for vehicles using an experimental optimization design technique on a wind tunnel test rig of the radiator. The authors have developed the regression equations of heat dissipation rate, coolant pressure drop and air pressure drop. The influences of the air velocity, inlet coolant temperature and volume flow rate of coolant on heat dissipation rate, coolant pressure drop and air pressure drop have been discussed in detail by means of the numerical analyses. The results published in this paper provide a basis for the theoretical analysis of heat performances and structural refinement of the tube-and-fin radiator.

Coupling of CFD and shape optimization for radiator design, a paper by Sridhar Maddipatla, [4] forms the benchmark for the present work. This paper presents a method to design automobile radiator by coupling CFD with a shape optimization algorithm on a simplified 2D model. It includes automated mesh generation using Gambit, CFD analysis using Fluent and an in-house C-code implementing a numerical shape optimization algorithm are discussed. All of the flow simulations reported in this paper were performed using the classical simple algorithm with a k- $\epsilon$  turbulent model and second order upwind scheme.

The study is aimed to analyze the fluid flow characteristics in a commercially existing radiator and understand the flow phenomenon to establish better design. It involves calculating the overall pressure drop and mass flow rate distribution of the coolant and air in and around the single tube arrangement of an automotive radiator. The fluid flow simulation is conducted using commercial software ANSYS 12.1.

## II. NUMERICAL EXPERIMENTATION

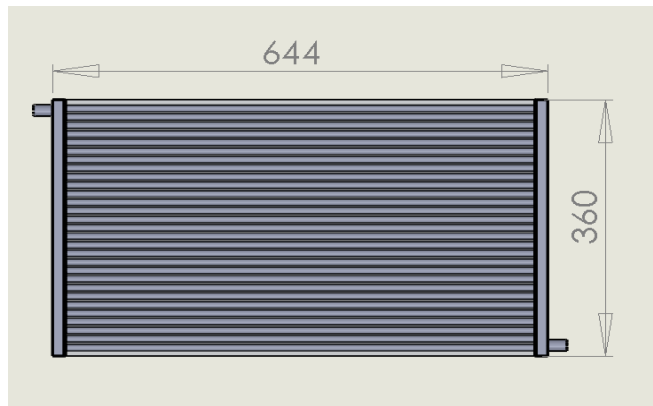
The Numerical experimentation of the work carried out involved reverse engineering of a commercial automotive radiator for the required fluid domain, discretizing the fluid domain, simulation of the fluid flow and heat transfer at steady state and post processing the results and drawing suitable conclusions. The radiator of the **Tata indigo diesel car** shown in figure (1) is chosen for the analysis to bring in the practicality to the study. The details of the geometry of the radiator were obtained by the process of reverse engineering. The dimensions of individual components of the radiator were measured using suitable measuring instruments. The measurements obtained were used to generate the CAD model in SOLIDWORK BENCH 9.1. Radiator is considered as a Shell and Tube Type Heat Exchanger. The specifications of this experimental radiator are mentioned below.

### Shell Side Data :-

Media: - Air  
 Temperature: - 30°C  
 Inlet Velocity: - 30 Km/h (Vehicle Speed)  
 Outlet Pressure: - 1.01325 bar



Experimental Radiator (**Tata indigo diesel car**)



Overall Dimensions of **Tata indigo car** Radiator

### Tube Side Data:-

Diameter of Tube: - 7 mm  
 No. of Tubes: - 29  
 Media: - Water + Ethanol  
 Temperature (Engine):- 70 °C  
 Inlet Velocity: - 2m/s  
 Outlet Pressure: - 1.01325 bar

Experimental Results of the radiator achieved is stated below.

Sr No.	Velocity of Car kmph	Velocity of Car m/s	Engine Temperature Tube Side Inlet	Tube Side Outlet Temperature	Shell Side inlet temperature	Shell Side Outlet temperature
1	30	8.33	70	51.23	30	33.14
		8.33	70	50.24	29	33.86
		8.33	70	49.86	28	34.56
		8.33	70	48.86	27	34.96
		8.33	70	48.02	26	35.12
		8.33	70	47.86	25	35.98
		8.33	70	48.86	27	34.96
2	40	11.11	70	54.15	30	31.15
		11.11	70	52.64	29	30.52
		11.11	70	51.24	28	29.85
		11.11	70	50.89	27	28.86
		11.11	70	49.86	26	27.85
		11.11	70	48.86	25	26.52
		11.11	70	48.56	25	26.52

Velocity of Car m/s	Engine Temperature Tube Side Inlet	Tube Side Outlet Temperature (Experimental)	Shell Side inlet temperature	Shell Side Outlet temperature (Experimental)	Heat Transfer Coefficient
8.33	70	51.23	30	33.14	3.85384718
8.33	70	50.24	29	33.86	5.96487176
8.33	70	49.86	28	34.56	8.05134953
8.33	70	48.86	27	34.96	9.76962535
8.33	70	48.02	26	35.12	11.1933396
8.33	70	47.86	25	35.98	13.4761917
11.11	70	54.15	30	31.15	1.88248593
11.11	70	52.64	29	30.52	2.48815532
11.11	70	51.24	28	29.85	3.02834694
11.11	70	50.89	27	28.86	3.04471638
11.11	70	49.86	26	27.85	3.02834694
11.11	70	48.56	25	26.52	2.48815532

Heat Transfer Calculation:-

Surface Area of Shell is  $A = h * l$   
 $= 0.36 * 0.61$   
 $= 0.2196 \text{ m}^2$

Mass Flow Rate is  $m = A * V * \rho$   
 $= 0.2196 * 8.33 * 1.1$   
 $= 2.012 \text{ Kg/s}$

Heat Transfer Rate is  
 $Q = m * C_p * \Delta T$   
 $= 2.012 * 1.0044 * (33.14 - 30)$   
 $= 3.85 \text{ KW}$

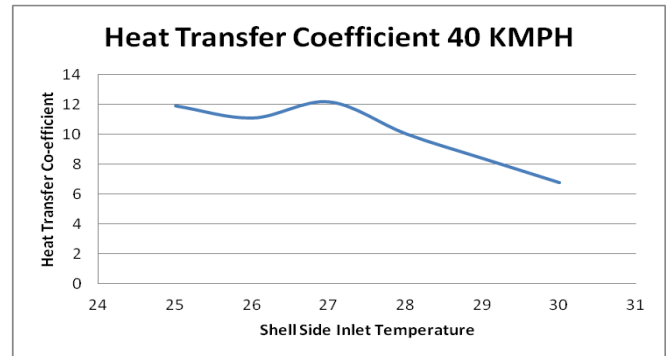
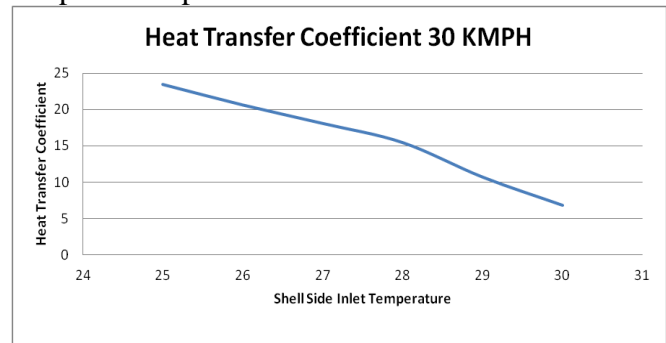
Overall Heat transfer Co-efficient is

$$U = \frac{Q}{A * (T_{outlet} - T_{mean})}$$

$$= \frac{3.85}{0.2196 * (33.14 - 30)}$$

$$= 5.583 \text{ KW/ m}^2 \text{ K}$$

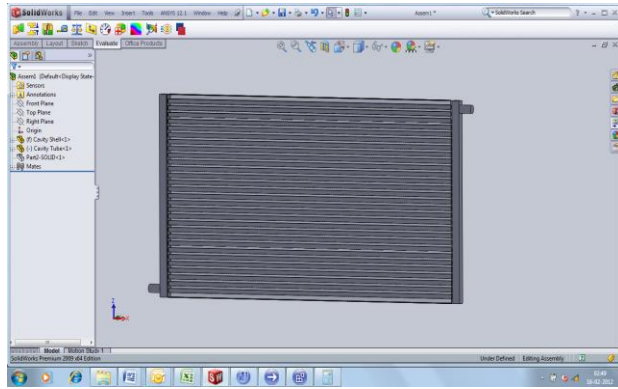
Graphs of Experimental Data:-



MODELLING OF RADIATOR

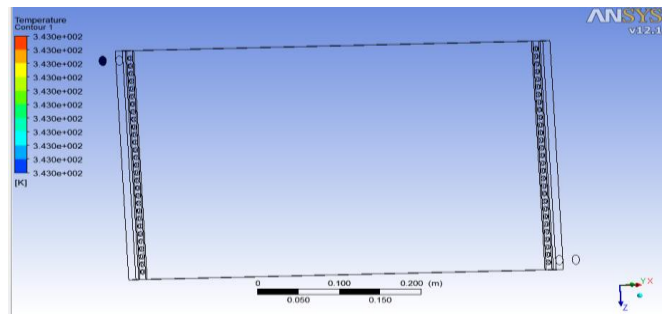
After performing simple calculation, the modeling has been performed on the Solid works 2009 version and then after the analysis work has been performed on the

ANSYS12.0version.

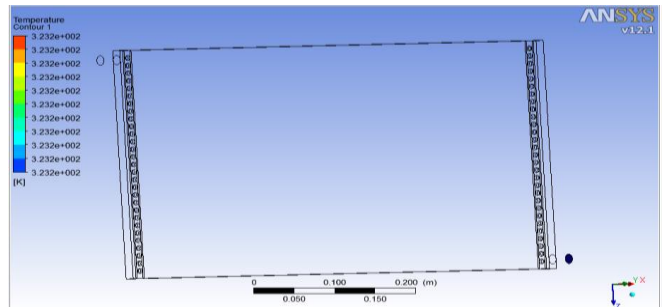


Results of Analysis

Tube Side Results



Inlet Temperature:- 343-273 :- 70°C

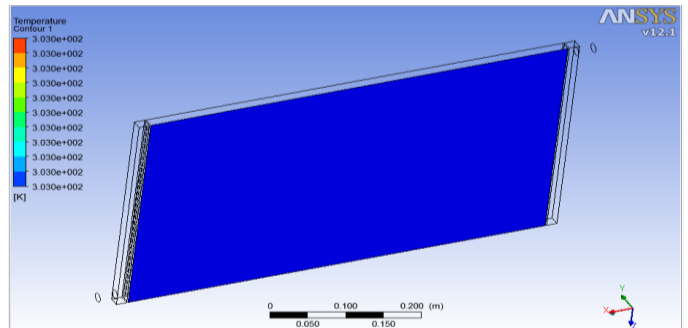


Outlet Temperature:- 323.23-273:- 50.23°C

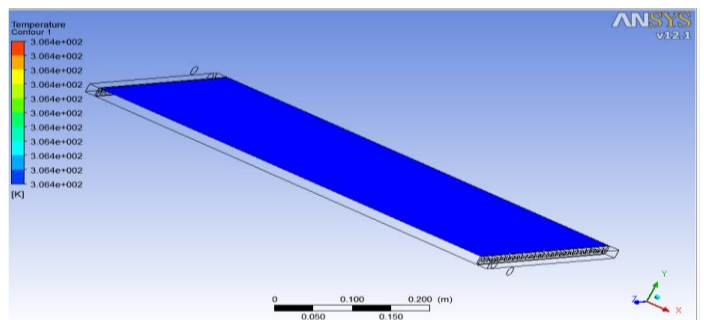
CFD ANALYSIS OF RADIATOR IN ANSYS CFX

The Cavity Pattern method is used for CFD Analysis of radiator in this study. First 3D model of Radiator is prepared in Solid works and then Cavity model. In cavity model, there is basically two Domains. D-1:- Water with addition of glycols & D- 2:- Air Domain Domain 1 Cool the engine block while Domain 2 Cool the Domain 1 Fluid. Now 3D Model is imported in ANSYS Workbench Mesh Module for Meshing. The input data and boundary conditions are chosen from the study of Changhua Lin and Jeffrey Saunders [5], in which the effect of varied air and coolant inlet temperatures on the specific dissipation is discussed. The air inlet velocity ( $V_{ai}$ ) is chosen to be 4.4m/s air at ambient temperature [4,5]. With reference to the specifications of the coolant pump used for the vehicle chosen, coolant inlet velocity ( $V_{ci}$ ) is 0.0063 kg/s in each tube. The properties of air and coolant were defined for standard conditions and kept constant throughout the analysis. As Per above Procedure, We have done 8 iteration for different Velocity and inlet temperature configuration which results are as below.

Shell Side Results:-



Inlet Temperature:- 303-273 :- 30°C



Outlet Temperature:- 306.35-273 :- 33.35°C

Sr No.	Velocity of Car kmph	Velocity of Car m/s	Engine Temperature	Tube Side Outlet Temperature	Shell Side inlet temperature	Shell Side Outlet temperature
1	30	8.33	70	50.23	30	33.35
		8.33	70	49.63	29	34.26
		8.33	70	48.52	28	35.6
		8.33	70	48.2	27	35.9
		8.33	70	47.2	26	36.15
		8.33	70	46.5	25	36.54
2	40	11.11	70	52.12	30	32.5
		11.11	70	51.56	29	32.1
		11.11	70	50.23	28	31.71
		11.11	70	49.63	27	31.5
		11.11	70	48.52	26	30.1
		11.11	70	47.8	25	29.4

CFD VALIDATION

To validate the CFD results, comparisons were drawn between obtained results and received experimental data which is given below.

Velocity of Car m/s	Engine Temperature Tube Side Inlet	Tube Side Outlet Temperature (Experimental)	Tube Side Outlet Temperature	Shell Side inlet temperature	Shell Side Outlet temperature (Experimental)	Shell Side Outlet temperature	Percentage variation Shell Side Temperature	Percentage Variation Tube Side Temperature
8.33	70	51.23	50.23	30	33.14	33.35	0.62968	1.95198
8.33	70	50.24	49.63	29	33.86	34.26	1.16754	1.21417
8.33	70	49.86	48.52	28	34.56	35.6	2.92134	2.68752
8.33	70	48.86	48.2	27	34.96	35.9	2.61838	1.35079
8.33	70	48.02	47.2	26	35.12	36.15	2.84923	1.70762
8.33	70	47.86	46.5	25	35.98	36.54	1.53256	2.84162
11.11	70	54.15	52.12	30	31.15	32.5	4.15384	3.74884
11.11	70	52.64	51.56	29	30.52	32.1	4.92211	2.05167
11.11	70	51.24	50.23	28	29.85	31.71	5.86565	1.97111
11.11	70	50.89	49.63	27	28.86	31.5	8.38095	2.47592
11.11	70	49.86	48.52	26	27.85	30.1	7.47508	2.68752
11.11	70	48.56	47.8	25	26.52	29.4	9.79591	1.56507

Above Results Table Shows that CFD Analysis Results Fairly matches with the experimental Results which show that CFD Analysis is a good tool for avoiding costly and time consuming experimental work.

EFFECT OF TUBE DIAMETER ON HEAT TRANSFER CO-EFFICIENT.

In this paper, the tube diameter of 14mm with nos. of tubes 13 and boundary condition were considered as below.

Shell Side:-

Inlet Temperature: - 30°C  
 Inlet Velocity: - 30 Kmph  
 Outlet Pressure: - 1.03125bar

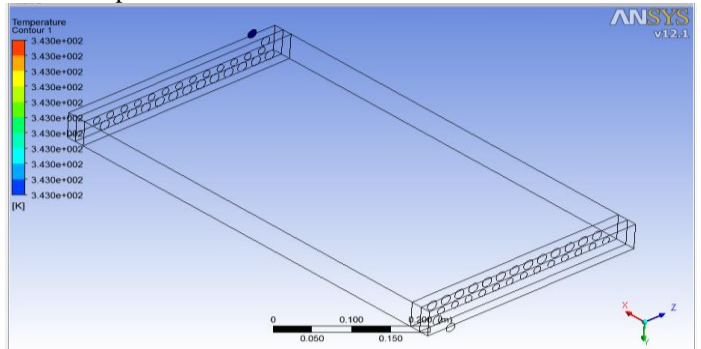
Tube Side:-

Inlet Temperature: - 70°C  
 Inlet Velocity: - 2 m/s  
 Outlet Pressure: - 1.01325 bar

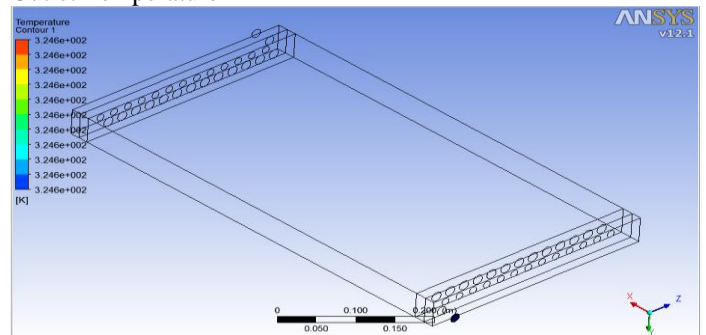
Result:-

Tube Side Results:-

Inlet Temperature

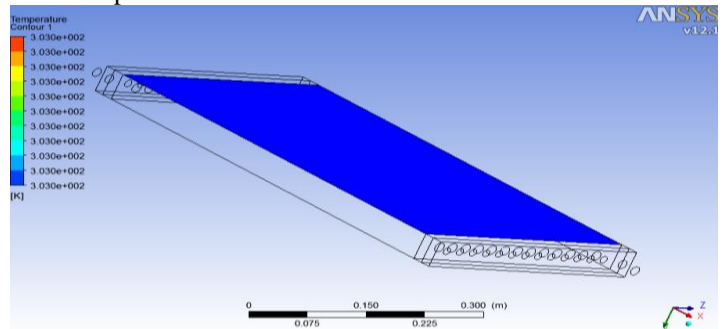


Outlet Temperature

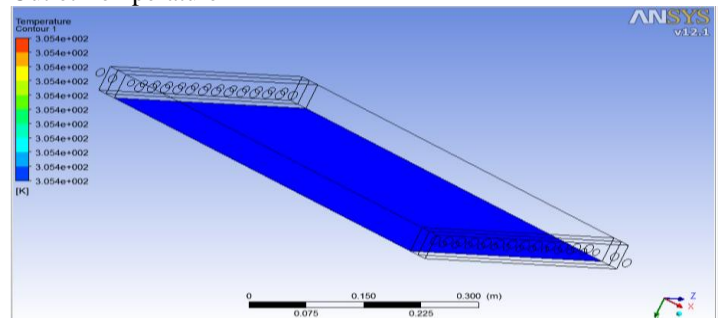


Shell Side Results:-

Inlet Temperature



Outlet Temperature



Similarly, the following results is obtained by the tube diameter of 5 mm with numbers of tubes 32 applying same boundary conditions.

Tube Diameter	Velocity of Car kmph	Velocity of Car m/s	Engine Temperature Tube Side Inlet	Tube Side Outlet Temperature	Shell Side inlet temperature	Shell Side Outlet temperature
14	30	8.33	70	51.56	30	32.36
05	30	8.33	70	50.95	30	32.9

### III. CONCLUSIONS

The fluid flow and heat transfer analysis of an automotive radiator is successfully carried out using numerical simulation built in commercial software ANSYS 12.1. Above Results Shows that as the diameter of tube is decreased, efficiency of tube decreases. So we can say that optimum efficiency is coming at the Diameter of 7 mm. The study forms a foundation for the fluid flow analysis of an automotive radiator. With the computational time and resources available, the results obtained were found to be satisfactory. However, to account for the variation of the inlet conditions with time as in practical cases, transient analysis can be done

### IV. REFERENCES

- [1] Hilde Van Der Vyer, Jaco Dirker and Jousoa P Meyer, 2003, "Validation of a CFD model of a three dimensional tube-in-tube heat exchanger", Third International Conference on CFD in the Minerals and Process Industry, CSIRO, Melborne, Australia. pp. 25-32.
- [2] A.Witry M.H. Al-Hajeri and Ali A. Bondac, 2003, "CFD analysis of fluid flow and heat transfer in patterned roll bonded aluminium radiator", 3<sup>rd</sup> International conference on CFD, CSIRO, Melborne, Australia, pp. 12-19.
- [3] J A Chen, D F Wang and L Z Zheng, 2001, "Experimental study of operating performance of a tube-and-fin radiator for vehicles", *Proceedings of Institution of Mechanical Engineers*, Republic of China, 215: pp. 2-8.
- [4] Sridhar Maddipatla, 2001, "Coupling of CFD and shape optimization for radiator design", Oakland University. Ph.D. thesis.
- [5] J.P.Holman, 2002, Heat transfer, Tata-McGraw-Hill Publications.
- [6] Incropera, F.P.; and DeWitt, D.P. (2002). *Fundamentals of heat and mass Transfer*. (5th Ed.), Wiley, New York.

# DYNAMIC SIMULATION AND STABILITY STUDY OF A PRESSURE REGULATOR

<sup>1</sup> M.R.PATEL, <sup>2</sup> PROF. KEDAR BHOJAK

<sup>1</sup>M.E.[Production Engineering] Student, Department Of Mechanical Engineering, L.D.R.P. Institute Of Technology & Research, Gandhinagar, Gujarat.

<sup>2</sup> Asst. Professor, Department Of Mechanical Engineering, L.D.R.P. Institute Of Technology & Research, Gandhinagar, Gujarat.

[milu.patel1988@gmail.com](mailto:milu.patel1988@gmail.com), [kedar.bhojak@oswalproject.com](mailto:kedar.bhojak@oswalproject.com)

**ABSTRACT:** With the increase of gas consumption and the expansion of the associated distribution network, a research program was set up to study the stability and to simulate the main characteristics of the dynamic behaviour of any type of pressure regulator. The modelling of a pressure regulating station is based on hydraulic, mechanical and valve models. The library of models is validated on one type of pressure regulator and simulations are in good agreement with measurements. The study showed that the operating conditions and installation requirements have the greatest influence on the stability of the pressure regulator. From measurements and simulations, the amplitudes of the downstream pressure are particularly sensitive to the size of the downstream volume and to upstream pressure.

**Keywords**— Regulator, stability, simulation, dynamic analysis, numerical method

## 1.1 INTRODUCTION

The many remarkable characteristics of pneumatic systems such as energy-saving, simple structure and operation, high efficiency and suitability for working in harsh environments, has led to their wide use for many years in different industries such as petroleum and aerospace engineering [1,2]. Particularly, gas pressure regulators are widely used in diverse applications to control the operational pressure of the gas. Gas regulators are devices that maintain constant output pressure regardless of the variations in the input pressure. They range from simple, single stage [3,4] to more complex, multi stage regulators [5,6], but the principle of operation is the same in all of them [7]. Limited research is published on pressure regulators, due to concerns over proprietary information. It should be pointed out that, for many years, the trial-and-error method of developing various kinds of pressure regulators was the only accepted way. Most probably, this was the result of the inherent difficulties in the modeling of these regulators. The difficulties encountered in making a dynamic analysis of the gas regulator are largely due to the nonlinearities of the equations describing the regulator. Tsai and Cassidy [8] introduced, probably for the first time, a systematic analysis, linear as well as nonlinear for a simple pressure regulator. Dustin presented an analog simulation of spring loaded, direct acting, single stage gas pressure regulator. The simulation was based on a pressure regulator used in a space power generator. In this investigation, the

effects of regulator design parameters on stability, transient response and steady state accuracy were examined. Kakulka et al. studied the piston pressuresensing unit regulator that had a conical poppet valve for regulating the gas flow. The dynamic effects of restrictive orifices and the upstream and downstream volumes were addressed in the modeling and analysis. Zafer et al. developed a comprehensive dynamic model for a gas pressure regulator in order to gain a better understanding of its behavior. They first modeled an existing regulator and used empirical data as necessary to identify parameter values for the model. Using a linearized version of the mentioned model, they investigated the effects of parameter variations using classical root-locus technique. Indeed, their main work was to identify the most influential system parameters on the stability of the system.

## NOMENCLATURE

$C_f$  viscous damping (kg/s)  
 $C_p$  specific heat at constant pressure (J/K/kg)  
 $D$  pipe diameter (m)  
 $E$  specific internal energy (J/kg)  
 $G$  gravity acceleration (m/s<sup>2</sup>)  
 $h$  specific internal enthalpy (J/kg)  
 $K$  spring stiffness (N/m)  
 $L$  pipe length (m)  
 $M$  actuator mass(kg)  
 $P$  pressure (Pa)  
 $P_a$  downstream pressure (Pa)

- $P_e$  upstream pressure (Pa)
- $PMC$  driving pressure (Pa)
- $PAM$  auxiliary pressure (Pa)
- $R$  gas constant (J/kg/K)
- $Q$  mass flow rate (kg/s)
- $U$  bulk mean velocity (m/s)
- $V$  downstream volume (m<sup>3</sup>)
- $Z$  compressibility coefficient (-)
- $u$  fluid velocity (m/s)
- $x$  position along the pipe (m)
- $X$  actuator position (m)
- $\beta$  coefficient of thermal expansion (K<sup>-1</sup>)
- $\phi$  conductive heat flux (W/m<sup>2</sup>)
- $\lambda$  pressure loss coefficient (-)
- $\mu$  dynamic viscosity (Pa.s)
- $\rho$  gas density (kg/m<sup>3</sup>)
- $\tau_{ij}$  shear stress (s<sup>-1</sup>)

### EXPERIMENTAL APPROACH

The first approach applied to study the conditions that lead to instabilities downstream of pressure regulating stations, consists in isolating one of the most common pressure regulator and install it on a test bench. The objective being to identify the configurations that enhance downstream pressure oscillations.

### Pressure regulators design

There are several class of pressure regulator, the role of all of them is to bring gas flow from a high pressure to a lower pressure. The gas expansion corresponds to an energy loss that is contain within a regulating action , the set variable can be the upstream pressure, the downstream pressure or the flow rate, only downstream pressure regulator are considered in this study.

A regulator is made basically of a control member, here a movable plug which is positioned in the flow path to restrict the flow. The control member is driven by an actuator : a diaphragm, dividing a casing into two chambers, providing the thrust to move the controlled member. One chamber is connected to downstream volume through a sensing line and the pressure induced force exerted on the diaphragm is balanced by the set value of the downstream pressure. Regulators work pneumatically with a power autonomy, they are divided into two main categories: direct acting and pilot controlled regulator. Basically, both are made of the same components, although the pilot controlled regulator can incorporates more components. For a direct acting regulator, the thrust is balanced by the load of a calibrated spring, whereas for a pilot controlled regulator, it is balanced by a controlled pressure: the auxiliary pressure (PAM) set in the upper casing(FIG 1)

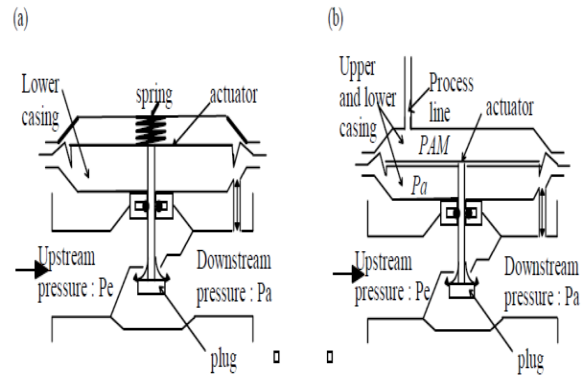


Figure 1. Schematic of direct acting regulator (a) and pilot controlled regulator (b)

The auxiliary pressure is supplied by expansion of upstream gas through a direct acting regulator: the pilot. In fact, prior to the pilot, another direct acting regulator: the pilot supply regulator prevents upstream disturbances by reducing the upstream pressure down to the pilot feeding pressure (PMC)(figure 2).

In brief, a pilot controlled regulator is then composed of three regulators: the main regulator, the pilot and the pilot supply regulator, it is also fitted out with several sensing lines and two valves, the first one: the pilot restrictor restrain the flow between the pilot regulator and the downstream pipe, the second one: the antipumping valve restrain the flow out of the lower chamber of the main regulator and is supposed helping to reduce the amplitude of oscillations.

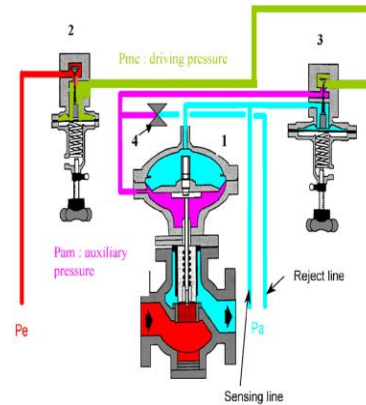


Fig. 2. Schematic of pilot controlled regulator  
(1: main regulator, 2: pilot supply regulator, 3: pilot, 4: creeper valve).

Thus, a pilot controlled regulator is composed of:

**controlled valve:** the movable part of the regulator which is positioned in the flow path to restrict the flow through the regulator;

**actuator:** the mechanism that makes the controlled valve move depending on the pressure in the two chambers;

**pilot:** its role is to compare the downstream pressure ( $P_a$ ) communicated to it by the sensing line. It



operates on the main regulator by using the auxiliary pressure (Pam) which makes the actuator move in the desired direction;

**pilot supply regulator:** its role is to make the regulation independent of the upstream pressure and to provide a constant driving pressure (Pmc);

**actuator case:** the house of the actuator. When the pressure in each chamber is different from atmospheric pressure, the chamber at the highest pressure is called the “motorized chamber”;

**reject and sensing lines:** the lines that connect impulse points to the regulator. The line with no internal flow is called the sensing line; the one with an internal flow rate is called the reject line.

## Experimental Set-Up

The regulator in test is a device with a DN 50 nominal diameter. The test bench set up in the OSWAL INDUSTRIES, GANDHINAGAR is able to reproduce the real working conditions of the regulator (flow-rate up to 10 000 m<sup>3</sup>(n)/h - inlet pressure up to Pe = 50 bar - outlet pressure controlled at Pa = 5 bar). The arrangement consists basically in a first regulator located upstream which control the inlet pressure, the regulator being tested, a piping system that enables to provide a variable capacity of the downstream buffer and a flow control valve that settles the load flow rate (figure 3).

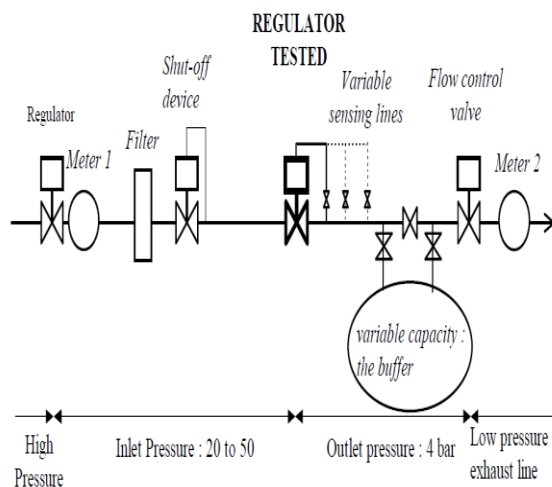


Figure 3. test bench

## Experimental Tests

Outlet pressure in downstream buffer and displacement of the flow control valve are registered after adjustment of several parameters: flow rate, inlet and outlet pressure, buffer size, length of the sensing lines, pilot feeding pressure, opening of pilot restrictor... The tests we carried out consisted in decreasing suddenly (in about 0.5s) the flow rate by 1000 m<sup>3</sup>(n)/h from the initial one in order to trigger pressure oscillations. The amplitude and frequency of oscillations depend on all these parameters. The influence of all of them has to be studied, but to limit

the number of tests, we applied experimental design approach.

## Experimental Design

The experimental design method requires a strategy which is different from usual approach that consists in changing only one parameter and keeping all the others constant. As a matter of fact the experimental design strategy is based on simultaneous variations of all the parameters. It is based on statistic and regression methods that does not require any knowledge concerning the physical phenomena involved in the system studied.

Concerning the pressure regulator, ten parameters were retained first [9,10]. But considering only two values per parameter would have expected 210 = 1024 experiments. In order to reduce this number, the experimental design used is based on fractional designs [7,8]. The first conclusion was that only five of them have a sensitive influence on the amplitude of pressure oscillations: the upstream pressure (Pe), the driving pressure (PMC), the downstream volume (V), the openings of the pilot restrictor and of the antipumping valve. In fact two experimental design have been performed for two flow rates : Q = 5000m<sup>3</sup>(n)/h and Q = 7000m<sup>3</sup>(n)/h.

In order to quantify the influence of these parameters and to interpolate, i.e. to calculate the amplitude at each point of the study field, another type of experimental design was used: composite designs. A similar approach has been then performed to conclude about simulation results.

## Modelling

The second part of the study, carried out in the same time as experiments, consists in developing an analytical model. It was considered that a better understanding of the physical phenomena involved is an appropriate method to define the operating conditions that maintain a suitable level of pressure for the regulator tested and then to extend the conclusions to any regulator. The physical model of a pilot controlled regulator requires several approximations that will have to be validated by comparisons of simulations and measurements.

### Generic equations of the hydraulic models

The method presented below presents the modelling of flow in pipes. The flow field is given by flow velocity u, pressure P, density ρ and temperature T. The model is based on equations of one-dimensional flow of a compressible, viscous, Newtonian fluid that are derived from the conservation of mass, momentum and energy completed with the equation of state.

### Conservation of mass

The difference between the mass flow rate entering and leaving a control volume induces changes in density. This leads to the equation:

$$\frac{d\rho}{dt} + \frac{\partial \rho u}{\partial x} = 0 \quad (1)$$

### Motion momentum

The equation of motion derives from Newton's law; the product of mass and acceleration is equal to the sum of external forces. We consider here only the pressure and friction forces, which act on the boundary of the fluid domain. The forces acting on the mass of fluid such as gravitational effects are not taken into account. We then obtain the following simplified form of the Navier–Stokes equations:

$$\int_{\Omega} \frac{\partial \rho \bar{u}}{\partial t} d\Omega + \int_{\Omega} (\overline{u \bar{n}}) \bar{u} \, ds = \int_{\Omega} -p \bar{n} \, ds + \int_{\Omega} \tau \bar{n} \, ds \quad (2)$$

This equation can also be written in the local form:

$$\frac{d(\rho u)}{dt} + \frac{d(\rho u^2)}{dx^2} = -\frac{dp}{dx} - \frac{\lambda}{2D} P l u u \quad (3)$$

The first two terms represent the inertia of the gas, the third, pressure forces and the fourth, friction forces.

However,

$$\frac{\lambda}{D} = \frac{\xi}{L} \text{ and } q = \rho u s \quad (4)$$

where  $\xi$  is estimated using the Idel'cik correlation [8]. Thus for an element for which cross sectional area is  $S$ , and characteristic length  $L$ , by using the two previous Eq. (4), the generic equation of motion **Type equation here.** conservation used for the hydraulic models is written in the form:

$$\frac{1}{s} \frac{dq}{dt} + \frac{1}{s} \frac{d(qu)}{dx} = \frac{dp}{dx} - \frac{\xi}{2L} \rho u l u \quad (5)$$

### Conservation of energy

The energy balance comes from the first law of thermodynamics. Change in total energy consists in changes in internal energy and in kinetic energy that correspond to the quantity of heat added and of the performed work that can be written in a local formulation

$$\rho \left( \frac{\partial e}{\partial t} + u_i \frac{\partial u_i}{\partial t} \right) = \frac{\partial p u_i}{\partial x_i} + \frac{\partial \tau_{ij} u_{ij}}{\partial x_i} - \frac{\partial \psi_i}{\partial x_i} \quad (6)$$

This formulation completed with the balance of kinetic energy (the product of Eq. (2) by the velocity) yields

$$\rho \left( \frac{\partial e}{\partial t} \right) = -u_i \frac{\partial p}{\partial x_i} + \tau_{ij} \frac{\partial u_j}{\partial x_i} - \frac{\partial \psi_i}{\partial x_i} \quad (7)$$

Last introducing the enthalpy defined by

$$e = h + \frac{p}{\rho} \text{ and } dh = de - d \left( \frac{p}{\rho} \right) = C_p dT + (1 - T\beta) dP \quad (8)$$

where  $\beta$  represent the coefficient of thermal expansion

$$\beta = -1/\rho \left( \frac{d\rho}{dT} \right) \quad (9)$$

The general energy equation can be written

$$\rho C_p \left( \frac{\partial T}{\partial t} + u_i \frac{\partial T}{\partial x_i} \right) = -\frac{\partial \psi_i}{\partial x_i} + T\beta \frac{dp}{dt} + \tau_{ij} \frac{\partial u_i}{\partial x_j} \quad (10)$$

The last term represents the contribution of viscous friction that has been neglected. In our first approach,

the walls are adiabatic and the only conductive heat flux comes from both ends of the pipe. Thus, for one-dimensional flow in an adiabatic pipe, the integral formulation of Eq. (10) is

$$\rho C_p \left( \frac{\partial T}{\partial t} + U \frac{\partial T}{\partial x_i} \right) = -S(\psi_{out} - \psi_{in}) + T\beta \frac{dp}{dt} \quad (11)$$

### Equation of state

The above equations do not give a complete description of the motion of compressible gas. The relationship between pressure variations and changes in temperature and density needs to be set through an equation of state.

We use the usual form:

$$\frac{p}{\rho} = ZRT \quad (12)$$

where  $R$  is the gas constant and  $Z$  the compressibility coefficient. The well-known perfect-gas ( $Z = 1$ ) approximation may not be suitable for this application since the pressure varies quite widely between upstream and downstream. For that reason the equation of state derived by

Peng–Robinson [10] has been used:

$$z^3 - (1 - B)z^2 + (A - 3B^2 - 2B)z - (AB - B^2 - B^3) = 0 \quad (13)$$

With

$$A = \frac{\alpha a P}{RT^2} \text{ AND } B = \frac{bP}{RT} \text{ for } T = 0.7T_c$$

$$a = 0.45724R^2 T_c^2 / P_c^2$$

$$b = 0.0778R \frac{T_c}{P_c}$$

$$\alpha =$$

$$[1 + (0.37464 + 1.54226w - 0.26992w^2)(1 - \sqrt{T_r})]^2$$

$$W = -\log_{10} \left( \frac{p_0}{p_c} \right) - 1 \quad (\text{where } P_c, T_c \text{ are pressure}$$

and temperature of the critical point, respectively).

### Generic equations of mechanical models

There are two different mechanical models for the three regulators since the same model is used for the pilot and the pilot supply governor, which are direct acting regulators. In fact, the nature of the forces acting on the diaphragms of these regulators are identical: pressure force in the upper casing (the motorization chamber) and spring stiffness in the lower casing. The pre-stressed spring of the pilot governor allows regulation of the pilot feed pressure whereas the characteristic of the pilot's spring is modulated to adjust the downstream pressure. The approach is somewhat different for the main regulator that integrates pressure forces on both sides of the diaphragm. Taking into account pressure forces, spring stiffness and damping, the motion of the plug is governed by

$$Mx = \Delta P_s - K(x - x_0) - M_g - F_\mu \text{sign}(\dot{x})$$

DP denotes the pressure difference between both sides of the diaphragm, s the diaphragm's area, Fm the friction

force resulting from the relative motion between the actuator and an O-ring seal. This is assumed to be Coulomb friction.

### Generic equations of valve models

There are two different models to represent the flow through the valves based on the pressure on both sides of the valve. The flow conditions vary widely whether sonic conditions are reached or not. The modelling for both sonic and subsonic valves are based on classical results and specific assumptions for a global approach. The model is based on established isentropic flow of a perfect gas.

Following Liepmann and Roshko [11], the energy equation leads to:

$$\frac{1}{2u^2} = \frac{\gamma}{\gamma-1} \frac{P_2}{\rho_1} \left(1 - \frac{P_2}{P_1}\right)^{\frac{\gamma-1}{\gamma}} \quad (14)$$

So assuming isentropic flow

$$\frac{P_2}{P_1} = \left(\frac{\rho_2}{\rho_1}\right)^\gamma = \left(\frac{T_2}{T_1}\right)^{\frac{\gamma}{\gamma-1}} \quad (15)$$

and considering that the pressure loss is quite small  $P_2/P_1 = 1 + \epsilon$ , Eq. (14) leads to

$$q = K_G \sqrt{P_1(P_1 - P_2)} \quad \text{with} \quad K_G = S \sqrt{2 \frac{\rho}{P} \frac{T}{T_1}} \quad (16)$$

where S denotes the flow area at the plug; P\*, T\*,  $\rho$ , respectively, stand for pressure, temperature and density at

the reference conditions. Sonic conditions are reached at the plug if the pressure downstream is lower than the critical pressure Pk [11] defined by

$$\frac{P_K}{P_1} = \left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}} \quad (17)$$

Eq. (14) can then be rewritten for sonic flow:

$$q = K_G' P_1 \quad \text{with} \quad K_G' = S \left(\frac{2}{\gamma+1}\right)^{\frac{1}{\gamma-1}} \sqrt{\frac{2\gamma}{\gamma+1} \frac{\rho}{P} \frac{T}{T_1}} \quad \text{and} \quad q = \rho_1 u S \quad (18)$$

The values of the parameters KG and  $K_G'$  depend significantly on the shape of the plug which is usually quite complex. They had to be determined for each valve by experiment. In practice, for natural gas it is usually

considered that  $K_G' \approx K_G/2$ .

### The Numerical Method

In this experiment this modelling was performed with Allan.@Simulation [17], a general software program designed for the modelling and simulation of technical and dynamic systems. It is a useful tool that serves design and analysis purposes. It has been built to enable engineers to work in their natural language using differential algebraic equations and technical diagrams.

When facing a complex system, the usual research engineer's method consists of breaking down the system into a set of elementary subsystems reduced to their essential behaviour, which are basically functional, technical or phenomenological.

Following this approach is one important concern for Allan.@Simulation as well as reducing the tasks related to computer programming or equation solving methods. In this aim, a box and string approach has been retained. The boxes, which are elementary models, contain the equations written in a very natural way. To be of assistance to the engineer in his design research, the equations must be able to handle both continuity (physical variables) and discontinuity (flags and booleans). A Fortran model can easily be incorporated in an assembly of Allan.@Simulation models. The equations used in the elementary physical models must describe their behavioural laws without conditioning how they will be used which can only be determined by implicit formulation.

The elementary models can be linked with others through strings, setting up a hierarchical ordered model that can be in turn connected with other composed or elementary models. The links, that can be oriented or non-oriented, hold one or several variables.

The algebraic-differential equations corresponding to the global model were solved by NEPTUNIX [18], the specific feature of which is the possibility of formulating equations implicitly and dealing with discontinuities.

Allan.@Simulation only deals with algebraic differential equations and not with partial differential equations as those derived in the set of equation (1),(4) and (11) which would have required a Computed Fluid Dynamic (CFD) solver to be properly treated. However, a CFD approach does not enable to represent the whole system specially as far as the motions of actuators are concerned. To enable the treatment of equations (1),(4) and (11) with Allan.@Simulation the pipes lengths have been cut into pieces using the following approximations for the gradients :

$$\frac{df}{dx} = \frac{f_2 - f_1}{dx} \quad (19)$$

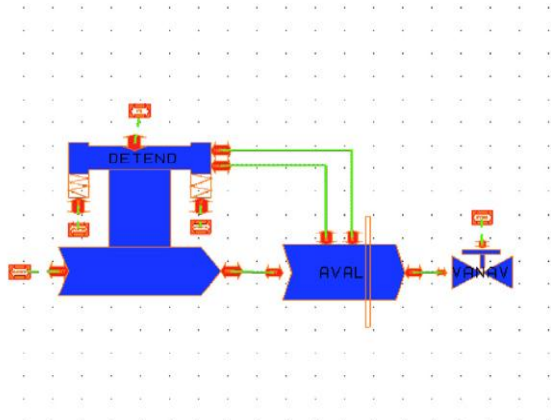
Where dx represents the length of one piece and  $f_i$  the value of function f at each end.

This approach does not lead to a very accurate spatial distribution, but is quite satisfying to take into account the main fluid dynamic phenomena involved in gas pressure regulators.

### The Global Model

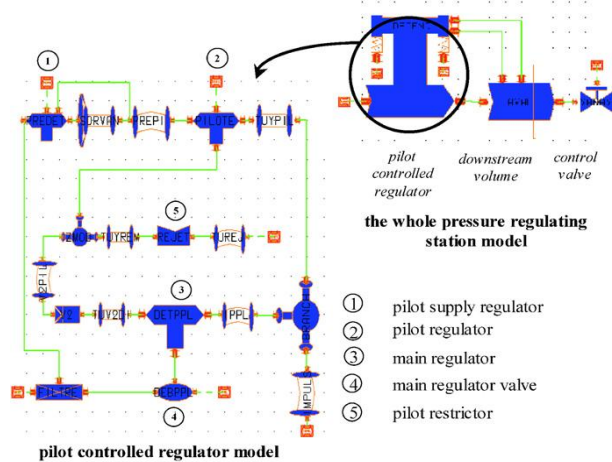
The global model is an assembly of three functional models that represent : the downstream valve, the downstream volume (the buffer) and the regulator itself (figure 4). The command statements are upstream pressure and temperature, setting of

springs of the pilot and supply regulator last the opening of the downstream valve.



**Fig 4 Schematic of global model**

The downstream valve is an elementary model that estimates the flow rate as a function of the valve opening and the pressure in the buffer. The buffer is an assembly of elementary models: pipes, junction points to connect to process lines. Last, the regulator model is a more complicated hierarchical model gathering models for the pilot, the pilot supply governor, the main regulator, pipes, chambers and valves: two different sonic valves for the main regulator and the pilot feeding regulator, and three subsonic valves: one for the pilot restrictor, one for the pilot regulator and the anti-pumping valve (figure 5).



**Fig 5 pilot controlled regulator model**

**SIMULATIONS AND MEASUREMENTS**  
**Experimental Observations**

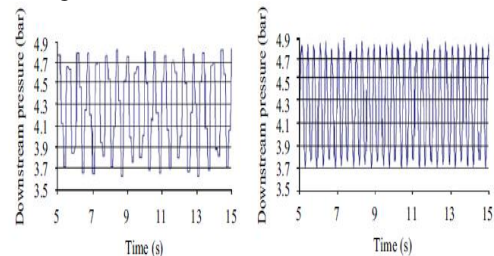
It turned out from first experimental results that the amplitudes of the oscillations vary widely with some parameters specially with upstream pressure, they depend also on the flow rate. An increase in flow rate tends to reduce the

amplitudes of the oscillations, while enhancing upstream pressure amplify the pressure instability downstream of the regulating station. This can be easily correlated with the position of the plug, as a matter of fact, the plug is almost closed in both cases when upstream pressure is high or when the flow rate is low. Another very important feature is the size of the gas buffer downstream of the regulating station, increasing the buffer's length leads to smaller pressure oscillations. The frequency of oscillations vary between 0.4 Hz. for long buffer and small upstream pressure up to 2.5 Hz for small buffer and high up-stream pressure. It turned out that the frequency vary mainly with the buffer size while it is not very sensitive to flow rate.

**Comparisons Between Experiments And Simulations**

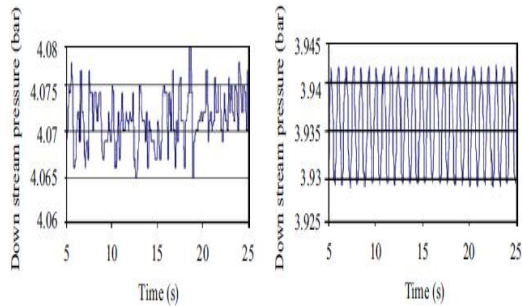
First, initializing calculations have been performed to get stationary conditions, then following the same method as for experiments, a perturbation on flow rate is done to generate oscillations.

In order to validate the model of the regulator, the numerical results are compared to measurements on test bench. It is generally noted that the numerical model reproduces with a satisfying level of accuracy the oscillations observed on test bench. The simulated amplitudes as well as the frequency are in a good agreement with measurements specially for small downstream volumes (figure 6).



**Fig 6 Simulations (---) and measurements (---) of downstream pressure for:  $V = 0.04m^3$ ,  $P_e = 20$  bar,  $q = 7000m^3(n)/h$**

The discrepancies between measurements and calculations become more important for high downstream volumes. The results presented on figure 7 have been obtained with  $V_b = 2.5m^3$  that corresponds to a length of about 80m. The differences may be due to coarse approximations concerning the representation of flow in long pipes. As discussed earlier, such a length would have required a more accurate approach to represent with a good level of confidence the compressibility or unsteady effects which is not coherent with the global system approach presented here. We considered a rough spatial discretisation, a series of five chaining pipes with linear gradients of pressure and this assumption is surely not realistic for such a long pipe.



**Fig 7 Simulations (---) and measurements (---) of downstream pressure for:  $V = 2.5m^3$ ,  $P_e = 20$  bar,  $q = 5000m^3(n)/h$**

However, it turned out from these last calculations that the model is able to reproduce the main tendencies as far as the influences of buffer size is concerned. Actually, both simulations and measurements indicate that amplitudes and frequencies of oscillations are reduced for highest downstream volume, the frequency simulated is generally higher than the frequency measured on test bench.

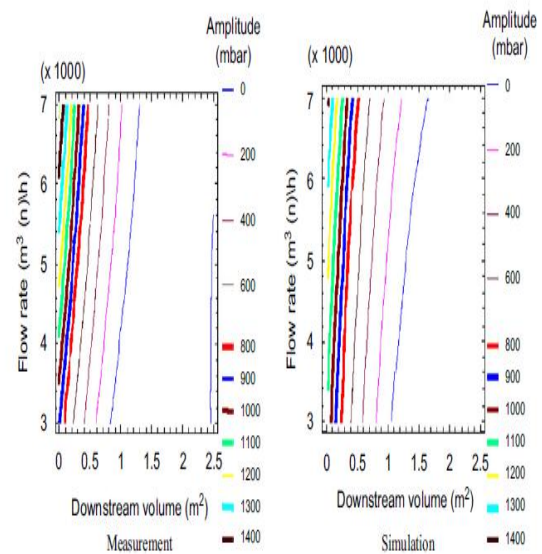
#### Qualitative validation

It turned out from a complete study that the amplitudes of oscillations vary widely with some parameters especially with downstream volume. They also depend on the flow rate and on the opening of the creeper valve.

Both simulations and measurements indicate that amplitudes and frequencies of oscillations are reduced for higher downstream volumes. According to the results in Fig. 8, a decrease in flow rate tends to reduce the amplitudes of oscillations. These results give a global illustration of the influence of downstream volume, flow rate on the amplitudes, and confirm that the most sensitive parameter is the size of the downstream buffer.

Note: The unit of the flow rate is in  $m^3(n)/h$ . It calculated in normal conditions: at  $P(n)$  of 1.01325 bar and  $T(n)$  of 273.15 K.

Another very important feature is the opening of the creeper valve. Three different openings of the creeper valve have been tested: 0.25; 0.5 and 1 (turn). Increasing the opening of the creeper valve for a small size of the downstream volume ( $V \frac{1}{4} 0.04m^3$ ) reduces the amplitudes considerably. The other parameters considered have less influence on the regulator stability. For example lengths of sensing lines, diameter of the downstream volume and driving pressure  $P_{mc}$ . Measurements and simulations have confirmed these features.



**Fig 8 Amplitudes of oscillations of downstream pressure obtained by calculations and measurement. For:  $P_e = 20$  bar,  $P_{mc} = 7$  bar and opening of creeper valve = 0.75.**

#### Conclusion

The final goal of the study presented in this paper is to improve the process control of pressure regulators. It is to define the operating conditions that maintain a constant pressure with oscillations within a given tolerance. For that purpose, numerical approaches have been performed.

The agreement between calculations and measurements indicates that the dynamic behaviour of a gas pressure regulator can be modelled mathematically and simulated with good accuracy.

It is generally noted that the amplitude of oscillations increases dramatically for small volumes ( $V \frac{1}{4} 0.04m^3$ ). Other parameters such as upstream pressure and flow rates have a significant influence on the oscillation of the downstream pressure, but their importance depends on the type of the pressure regulator.

In contrast, other parameters such as opening of creeper valve or antipumping valve, length of sensing line or driving pressure have only a small influence on the regulator stability. Nevertheless, from a quantitative point of view, some differences between simulations and measurements still remain in certain configurations. Some improvements will have to be brought to the model to represent with a higher level of accuracy the gas flow in the downstream volume. Some more work is underway to take into account the fluid forces acting on the plug.

It is believed that these computer models are accurate enough to be used to improve the design and the performance of regulators. It is believed that this work will find some applications in a larger domain or for other regulators to define

working conditions or to design full pressure regulating stations.

## References

- [1] J. Pu, P. R. Moore and C. B. Wong, "Smart Components-based Servo-pneumatic Actuation Systems", *Microprocessors and Microsystems*, Vol. 24, pp 113-119, (2000).
- [2] H. E. Merrit, "Hydraulic Control Systems", John Wiley & Sons, New York, (1967).
- [3] R. Mooney, "Pilot-Loaded Regulator: What You Need to Know", *Gas Industries*, pp 31-33, (1989).
- [4] R. Brasilow, "The Basics of Gas Regulators", *Welding Design and Fabrication*, pp 61-65, (1989).
- [5] A. Krigman, "Guide to Selecting Pressure Regulators", InTech, pp 51-65, (1984).
- [6] E. Gill, "Air-loaded Regulators", *The Smart Control Valve Alternative*, InTech, pp 21-22, (1990).
- [7] S.J. Bailey, "Pressure controls 1987: sensing art challenges old technologies", *Control Engineering*, pp 80-85, (1987).
- [8] Idel'cik, *Spravochnik po guidravlicheskim soprotivleniam* Moscow, Gosenergoizdat (translation in French: 1986 "memento des pertes de charges" Eyrolles), 1960.
- [9] Taine J, Petit JP, *Transfert de chaleur; Thermique; Ecoulement de fluide* Dunod, 1989.
- [10] Peng DY, Robinson DB. A new two-constant equation of state. *Ind Eng Chem Fund* 1976;15(1):59-64.
- [11] Liepmann HW, Roshko A. *Elements of gas dynamics*. New York: Wiley; 1962.
- [12] Jeandel A, Favret F, Lapenu L, Lariviere E, ALLAN. Simulation, a general tool for model description and simulation. Proceedings, IBPSA conference, Adelaide, 1993.
- [13] Nakhle M, NEPTUNIX an efficient tool for large size systems simulations. In second international conference on system simulation. Lieges, Belgium, December 3, 1991.
- [14] Delenne B., Mode L., Blaudez M., (1999), "Modelling and simulation of a gas pressure regulator" European Simulation Symposium, Erlangen Germany
- [15] Delenne B., Mode L (2000), "Modelling and simulation of pressure oscillations in a gas pressure regulator" Proceedings ASME 2000 FPST- vol 7-IMECE Orlando
- [16] ATG (1986) "Analysis of surging phenomena"
- [17] Adou K.J (1989) Etude du fonctionnement dynamique des régulateurs-detendeurs industriels de gaz PhD Thesis- Université Paris VI
- [18] Nakhle M (1991) "NEPTUNIX an efficient tool for large size systems simulations" In 2nd International Conference on System Simulation (Lieges, BELGIUM, Dec & 3)
- [19] Adou K.J. and J.P. Guiraud (07/1990) - "Modelling and non linear stability of piloted gas regulator" – European Journal of Mechanics B/Fluids
- [20] Favret F., Jemmali M.; Cornil J.P., Deneuve F., Guiraud J.P. (1990) "Stability of Distribution Network Governors" - R&D Forum Osaka Gas
- [21] Earney H.W (1996) "Causes and cures of regulator instability" Fisher controls international, Inc
- [22] Carmichael C (1996) "Gas regulators work by equalizing opposing forces" Pipe Line & Gas Industry pp 29-33

# Influence of off-axis loading on the mechanical properties of composite materials

A. Piyush S. Jain, B. A. A. Shaikh

**Abstract:** In last few decades, composites materials dominate aerospace, automotive, construction, leisure and sporting industries. Mechanical properties as Modulus, tensile strength, mode of failure, fracture toughness etc is not only dependent upon properties of resin and fibers but also dependent upon loading direction, fiber volume fraction and interfacial properties. Composites properties vary with the variation of off-axis loading. In this paper, effect of off-axis loading on the mechanical properties of composite materials are analyzed

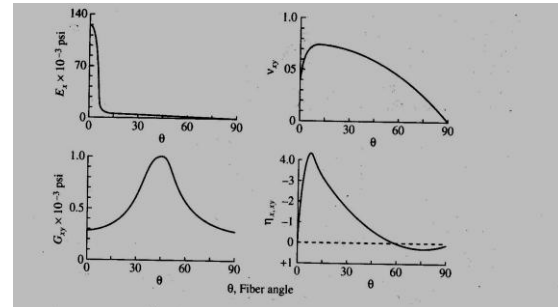
**Index terms:** off-axis loading, mechanical properties, composites, fibers

## I. INTRODUCTION

Polymer based composites have gained substantial interest over the last years due to very high strength-to-weight and stiffness-to-weight ratios. These properties led to the application of composite materials in a lot of major industries with the aerospace and auto industry taking the larger part. [1-2]. It is required to enhance the properties of matrix material so that overall properties of composites improved. It is found that some properties of composite materials depend upon various factors as properties of constituents, fiber volume fraction, fiber orientation and interfacial properties. When composite materials are prepared, the reinforcement is likely to be oriented in the load direction either intentionally or unintentionally. However, if the load direction is variable, and not parallel to the fibers, it becomes important to predict mechanical behavior of composite materials.

## II. Off-axis loading and mechanical properties

Halpin analysed the variation in engineering constant with fiber orientations. The results of Halpin are summarized in figure 1.



**Figure: 1. Variation of Engineering constant with fiber orientation [1]**

### A. Modulus and strength

L.E.Neilson & P.E. Chen [3] describes the prediction of modulus for the condition of fiber reinforcement based on classical theory which is given as

$$\frac{E_1}{E_\theta} = \cos^4 \theta + \frac{E_1}{E_2} \sin^4 \theta + \left( \frac{E_1}{G} - 2\nu_{12} \right) \cos^2 \theta \sin^2 \theta$$

On the same basis the model proposed by Horio & Onogi [4] observed to be more accurate considering the availability of experimental values of longitudinal and transverse values of composite material as given below.

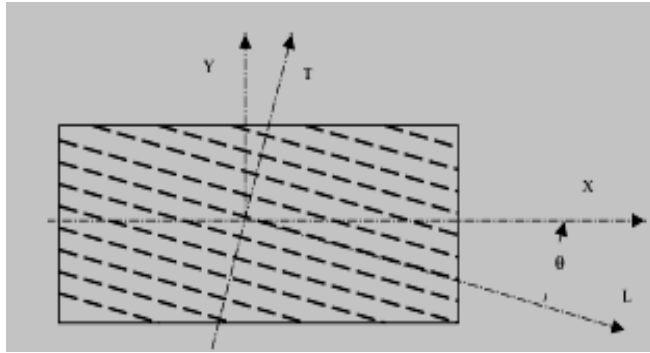
$$E_\theta = \frac{E_1 E_2}{E_1 \sin^2 \theta + E_2 \cos^2 \theta}$$

$$\frac{E_\theta}{E_1} = \left( \frac{E_2}{E_1} \right)^{\frac{1}{2}}$$

Baley [5] studied the tensile behaviour and analysis of the tensile stiffness increase of flax fiber. By using micromechanical equations, he estimated the Young's modulus of a flax fiber by taking into account the orientation of the fibrils during a tensile test.

A. Research scholar with S.V.N.I.T., Surat; (email:psjain45@gmail.com)

B. Associate professor with Mechanical Department, S.V.N.I.T., Surat; (email: aas@med.svnit.ac.in)



**Figure 2. Effect of fiber orientation on modulus [5]**

A micro mechanical model is proposed associating modulus with fiber orientation considering flax fiber as discontinuous but well aligned as given

$$E_x = \frac{1}{\frac{C^4}{E_L} + \frac{S^4}{E_T} + 2C^2S^2 \left( \frac{1}{2G_{LT}} - \frac{\nu_{LT}}{E_L} \right)}$$

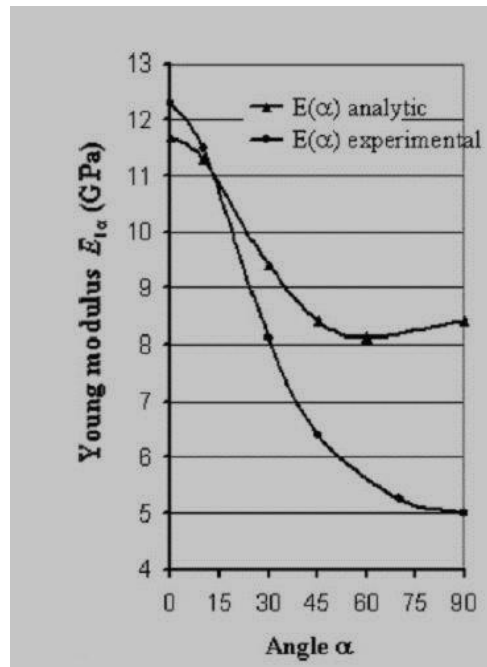
It is further simplified considering  $\nu_{LT}$  &  $\theta$  is very small.

$$E_x = \left( \frac{1}{E_L} + \frac{\theta^2}{G_{LT}} \right)^{-1}$$

Sami ben and Riddhi ben [6] evaluated the influence of fibers orientation and volume fiber fraction on the mechanical properties of the Alfa/Polyester composites. Figure 3 shows the experimental and theoretical evolution of the Young modulus  $E_L(\alpha)$  and the Poisson ratio  $\nu_{LT}(\alpha)$  with the fibers direction,  $\alpha$ . The theoretical evolution is given by the analytical expression as mentioned below:

$$G_{tt} = \frac{E_l E_t \cos^2 \alpha \cdot \sin^2 \alpha}{E_l \left[ 1 - E_l \left( \frac{\cos^4 \alpha}{E_l} + \frac{\sin^4 \alpha}{E_t} \right) \right] + 2\nu_{l,t} E_l \cos^2 \alpha \cdot \sin^2 \alpha}$$

$$\frac{1}{E_L(\alpha)} = \frac{\sin^4 \alpha}{E_t} + \frac{\cos^4 \alpha}{E_l} + \cos^2 \alpha \cdot \sin^2 \alpha \left( \frac{1}{G_{tt}} - \frac{2\nu_{l,t}}{E_l} \right)$$



**Figure 3. Experimental and analytical investigation of Young’s modulus with off-axis loading [6]**

U. C. Jindal [7] developed composites with bamboo fiber reinforced composites. These composites have been tested for tensile strength and stiffness. It is observed that ultimate and yield tensile strength increased with increasing volume fraction as shown in figure-4.

The ultimate tensile strength of bamboo fiber reinforced composites is more or less equal to mild steel though its density is almost one eighth as compare to mild steel.

The variation of young’s modulus along the direction of loading with different fiber orientation is also observed as shown in figure 5. It is observed that BFRP composite possess maximum stiffness along the fiber and minimum stiffness when fibers are orientated at an angle of 60. It is also observed that variation of modular ratio with fiber orientation is similar as the variation of modular ratio of another composite



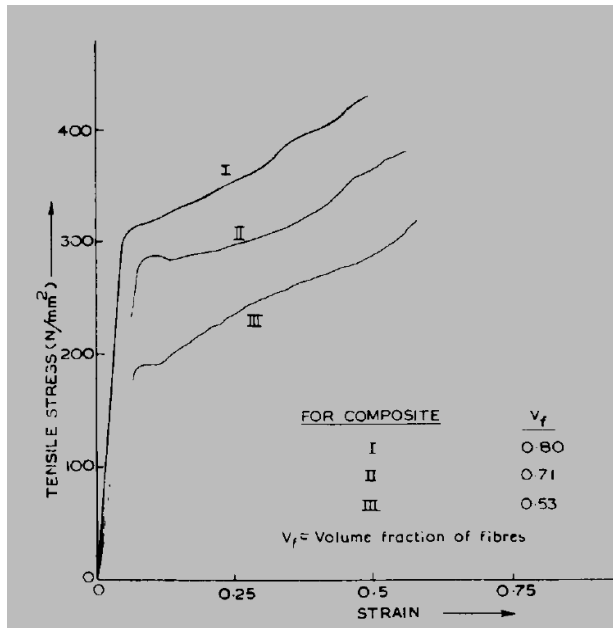


Figure: 4. Stress strain curves of BFRP Specimen with varying volume fraction [7]

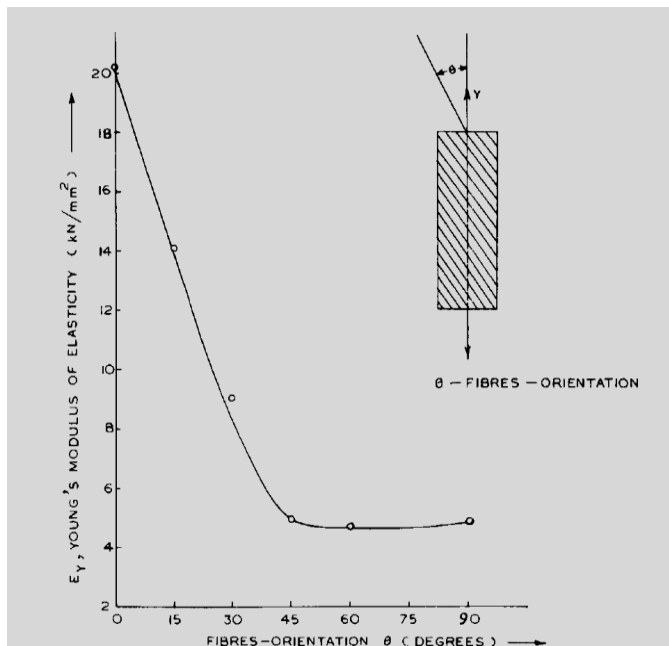


Figure: 5. Variation of Modulus with fiber orientation in BFRP Composites [7]

transverse modulus of elasticity. Variation of modular ratio is observed with different type of composites. It is observed that the behavior of BFRP Composites is similar to artificially developed composites.  $E_y/E_2$  curve for BFRP composite lies between GFRP composites and boron fiber reinforced composites.

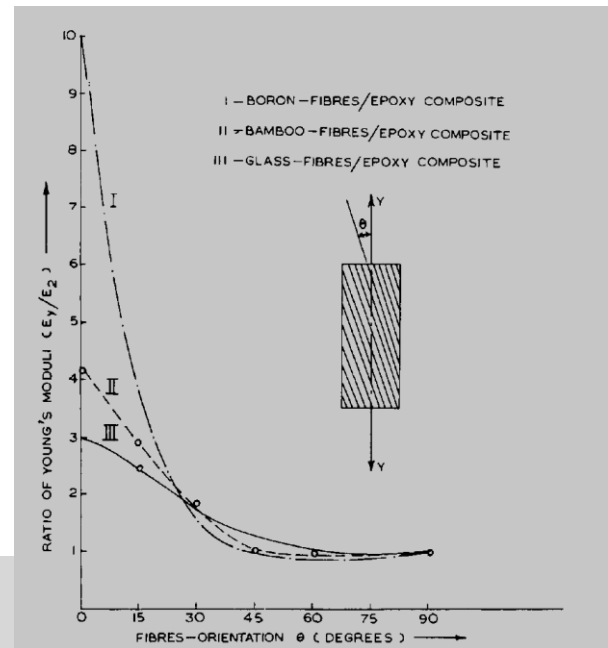


Figure: 6. Variation of Modulus with fiber orientation in various composites [7]

B. D. Agrawal and J. N. Narang [8] analyzed the behaviour of specimens of glass fiber-reinforced epoxy with various fiber orientations under tensile loading. Epoxy resin was selected as matrix material and continuous glass fibers are selected as reinforcement. Tensile tests were performed on the specimens with the fibers inclined at angles of 0, 10, 20, 30, 45, 60, 70, 80 and 90 degrees with the loading axis. At least five specimens were tested for each fiber orientation. The strength of the composites has been plotted as a function of the fiber orientation in Figure 7. It may be noted that the strength of the composites drops very sharply as the angle between the fibers and the loading axis increases.

The variation of modular ratio ( $E_y/E_2$ ) with different composites is shown in figure-6.  $E_y$  is modulus of elasticity along the direction of loading and  $E_2$  is

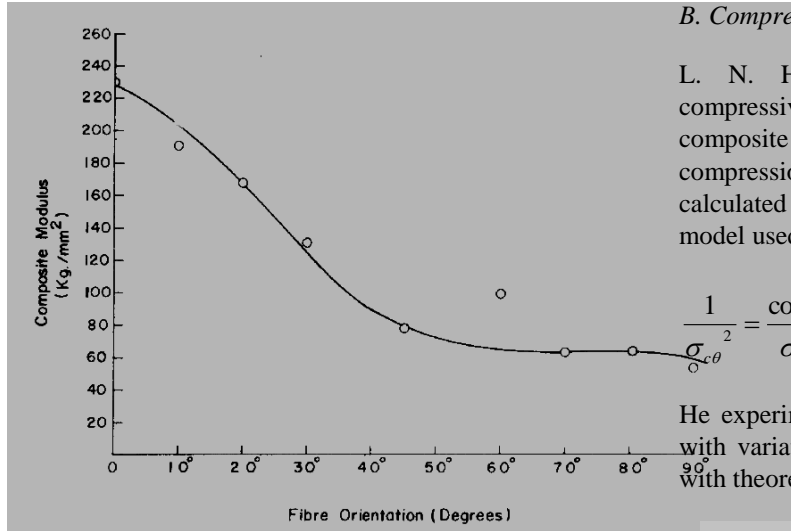


Figure: 7. Modulus of the composites as a function of fiber orientation [8]

The experimental results can be compared with the predictions of various failure theories for the composites as shown in figure 8. One of the more widely accepted theories is the maximum work theory which predicts the strength of the composites as follow:

$$\frac{1}{\sigma_{\theta}^2} = \frac{\cos^4 \theta}{X^2} + \left( \frac{1}{S^2} - \frac{1}{X^2} \right) \cos^2 \theta \sin^2 \theta + \frac{\sin^4 \theta}{Y^2}$$

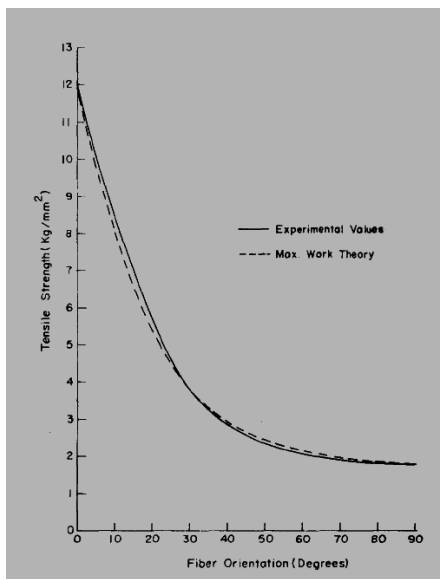


Figure: 8. Strength of the composites as a function of fiber orientation [8]

B. Compressive Strength

L. N. Hancox [9] studied the variation of compressive strength of a 60 vol % carbon fiber composite with the angle between the fiber and compression axis. The theoretical curve was calculated from the maximum work theory. The model used was

$$\frac{1}{\sigma_{c\theta}^2} = \frac{\cos^4 \theta}{\sigma_{cl}^2} + \left( \frac{1}{\tau^2} - \frac{1}{\sigma_{cl}^2} \right) \sin^2 \theta \cos^2 \theta + \frac{\sin^4 \theta}{\sigma_{ct}^2}$$

He experimentally studied the compressive strength with variation of compressive axis and compared it with theoretical model as shown in figure 9.

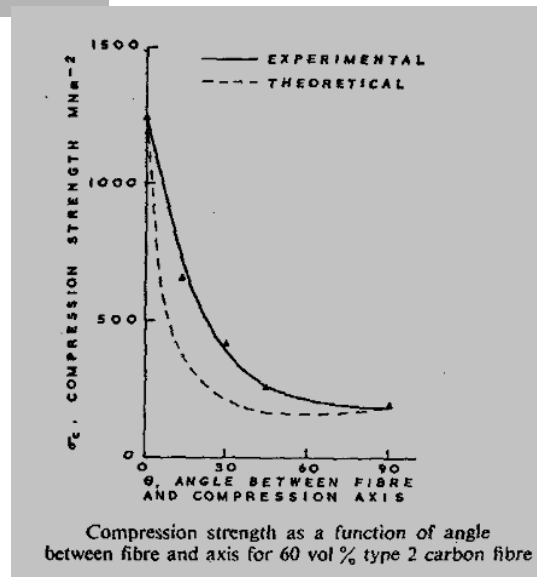


Figure: 9. Effect of orientation angle on compressive strength [9]

C. Fracture Toughness

G. S. Giare and R. R. Newcomb [10] studied the effect of slight orientation of fiber on fracture toughness of unidirectional fiber reinforced composites. Experimental investigations were performed on unidirectional glass fiber reinforced/epoxy laminates having fiber oriented at  $\theta = 0, \pm 1.2, \text{ and } \pm 2.4$  respectively in the presence of a crack parallel to the neutral axis of the beam. He observed that fracture toughness increased considerably with fiber orientation for  $\theta = \pm 1.2, \text{ and } \pm 2.4$  as compared to fiber orientation at  $\theta = 0$ . He found that slight fiber orientation increased the crack

propagating area, thus causing an increase in the fracture toughness of the composite. Slight fiber orientation had no effect on the longitudinal strength of the laminates. It is found that fracture toughness increase by approximately 100% by just orienting the fiber to  $\pm 2.4$ , as compared to unidirectional fiber reinforced composites ( $\theta = 0$ ) without decreasing the longitudinal strength of the composite.

### III. CONCLUSION

Effect of off-axis loading on the mechanical properties can be summarized as below:

Properties of composite materials depend upon various factors as constituent property, fiber content, fiber direction with respect to loading axis, adhesion characteristics between binder and reinforcement etc.

Effect of off-axis loading is also dependent upon fiber and binder material.

It is observed that tensile strength and modulus exhibit maximum properties in longitudinal direction. However least properties might be obtained at any direction of fiber (between) 60 to 90 with respect to loading axis. It is not necessary that minimum value of modulus and strength are obtained at transverse direction

Fracture toughness is very sensitive to fiber orientation.

### REFERENCES

- [1] F.L. Matthews and R.D. Rawlings, "Composite materials: Engineering and science", Woodhead publishing company England, 2001, pp. 1-10
- [2] Ronald F. Gibson, "Principles of Composite Materials Mechanics", Mc-Graw Hill International editions, 1994, pp. 1-5.
- [3] L.E. Neilson, & P.E. Chen, "Young's Modulus of Composite Filled with Randomly Oriented Fibers", Journal of Materials, Vol.3, No. 2, 1968, pp. 352-358.
- [4] M. Horio & S. Onogi, Journal of applied physics, vol. 22, 1951, pp. 971.
- [5] C. Baley, "Analysis of the Flax Fibres Tensile Behaviour and Analysis of the Tensile Stiffness Increase", Composites Part A, 2002, Vol 33, pp 939-948
- [6] Sami Ben Brahim and Ridha Ben Cheikh, "Influence of fibre orientation and volume fraction on the tensile properties of unidirectional Alfa-polyester composite", Composites Science and Technology, 2007, Vol 67, pp140-147
- [7] U. C. Jindal, "Development and Testing of Bamboo-Fibres Reinforced Plastic Composites", Journal of Composite Materials, 1986, Vol 20, pp 19-29
- [8] B. D. Agrawal and J. N. Narang, "Strength and Failure Mechanism of Anisotropic Composites", Fiber Science and Technology, 1977, Vol 10, pp37-52
- [9] N.L. Hancox, "The Compression Strength of Unidirectional Carbon Fiber Reinforced Plastic", Journal of material science Vol.10, 1975, pp. 234-242.
- [10] G. S. Giare & R. R. Newcomb, "Effect of Slight Orientation of Fiber on Fracture Toughness of Unidirectional Fiber Reinforced Composites in Mode II", Engineering Fracture Mechanics, Vol. 22, No 4, 1986, pp 667-679.



**Piyush S. Jain**

Education Quality Improvement Program funded by world bank as TEQIP Coordinator.

**Piyush S. Jain** completed Bachelor degree in Production Engineering from S V National Institute of Technology, Formerly known as S V Regional College of Engineering in 1998. He completed M. Tech in Mechanical Engineering in 2006 from S. V. National Institute of Technology, Surat (India). He is pursuing Ph D in Mechanical Engineering (Composite Material) from S. V. National Institute of Technology, Surat (India). His publication includes 1 International Journal, 1 National Journal, 3 International Conference and 1 National Conference.



**A. A. Shaikh**

**Dr. A. A. Shaikh** completed Bachelor degree and Post graduation in Mechanical Engineering from S V National Institute of Technology, Formerly known as S V Regional College of Engineering in 1987 and 1996 respectively. He completed Ph D in Mechanical Engineering (Composite Material) in 2005 from S. V. National Institute of Technology, Surat (India). He has guided 25 M. Tech. dissertation and at present 6 Ph D research work. His publication includes 18 journal papers and more than 40 papers in International and National Conference. He has actively organized short term training programs and workshops related to composite materials and CAD/CAE. He is actively involved in Technical

# A Review of Correlations of Condensation for Refrigerant R 22 & R 134a

A. Nimesh Gajjar, B. Nirav Joshi, C. Dr N M Bhatt

**Abstract**— Recent Vapour Compression Refrigeration (VCR) cycle based system mainly uses R134a as a substitute for R12 which was widely used previously in refrigerator, water cooler, car air conditioner etc. and R22 has many domestic and commercial applications like window air conditioner, central air conditioning systems, deep freezers etc. as a substitute for R11.

This work includes review of different correlation for condensation process for refrigerant HFC 134a and HCFC 22. It is also include performance of both refrigerants for different operating condition according to suitability of working equipments i.e. horizontal, vertical, smooth, rough, finned, grooved, size of parts etc. So, we can choose suitable refrigerant for our equipment.

**Index Terms**—Condensation, R-134a, R22, correlation, fin, channel, tubes, convection, Nusselt.

## Nomenclature

A - Cross sectional area ( $m^2$ )  
 Bo - Boiling number =  $q / G \cdot h_{fg}$   
 Co - Convection number =  $(1/X-1)^{0.8} (\rho_v / \rho_l)^{0.5}$   
 Cp - Specific heat (J/kg.K)  
 d - Diameter of tube (m)  
 E - Enhancement factor  
 F( ) - General function  
 F<sub>fl</sub> - Fluid-surface parameter  
 Fr - Froude number =  $G^2 / \rho_l^2 \cdot g \cdot d$   
 G - Mass flux ( $kg/m^2 \cdot s$ )  
 g - Acceleration due to gravity ( $m/s^2$ )  
 h - Heat transfer coefficient ( $W/m^2 \cdot K$ )  
 h<sub>fg</sub> - Latent heat of vaporization (J/kg)  
 ilv - Heat of Vaporization  
 k - Thermal conductivity (W/m.K)  
 L - Length of tube (m)  
 M - Molecular weight  
 Nu - Nusselt number  
 Pr - Prandtl number  
 p - Pressure  
 p<sub>sat</sub> - Saturation pressure (N/m<sup>2</sup>)  
 Q - Heat flux (W/m<sup>2</sup>)  
 q - Specific heat flux (W/m<sup>2</sup>.s)  
 Re - Reynolds number =  $G (1-X) d / \mu_l$   
 S - Suppression factor  
 T - Temperature (K)

U - Overall heat transfer coefficient ( $W/m^2 \cdot K$ )

X - Quality of refrigerant

X<sub>tt</sub> - Martinelli's coefficient

## Greek letters:

$\beta$  - Contact angle

c - critical

$\Delta$  - Increment

$\mu$  - Viscosity (N.s/m<sup>2</sup>)

$\rho$  - Density (kg/m<sup>3</sup>)

$\sigma$  - Surface tension (N/m)

$\psi$  - Dimensionless parameter (shah)

$\theta$  - Angle subtended from the top

## 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

Condensation and evaporation heat transfer of refrigerant occurs in many applications like refrigerator, air-conditioner and heat pumping system by phase change. So, for design of the heat exchangers, proper correlations must be selected for efficient design of them. Various different hydrodynamic conditions are encountered during refrigerant is condensed in the condenser so, this paper includes different methods of predict heat transfer coefficient for condenser. It is also include important feature that is often overlooked in comparing the correlations is the trends predicted by these correlations for heat transfer coefficient variation with quality. The objective of this paper is to choose simple and efficient form of generalize correlation for different geometry in which refrigerant is flowing because this study is justified by leading role that these heat exchangers now play in certain application.

A. AP, Mechanical Engg, Gandhinagar Institute of Technology  
 (e-mail: nimesh.gajjar@git.org.in)

B. AP, Mechanical Engg, Gandhinagar Institute of Technology  
 (e-mail: nirav.joshi@git.org.in)

C. Director, Gandhinagar Institute of Technology  
 (e-mail: director@git.org.in)

Due to quick destruction of the ozone layer in the earth atmosphere noted recently has been primarily related to the wide use of the CFC refrigerants. 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, thermo physical, flow and heat transfer properties. We realize that a much more detailed understanding of the flow condensation heat transfer of new refrigerants (R 134a) & R-22 is very important in the design of condensers used in many current refrigeration and air conditioning systems.

1. Condensation in smooth horizontal tubes:

In the case of condensation inside horizontal pipes, the phase-change process is dominated by vapour shear or gravity forces in such a way that the regime encountered depends on the force controlling the condensation process. Thus, if annular flow is the dominant process, it is because of the high vapour shear but if stratified, wavy or slug flows is found, it is due to gravity forces. It has been observed that in the annular flow regime, the heat transfer coefficient varies with mass velocity, G, vapor quality, x, and saturation temperature during condensation of pure fluids.

1.1. Cavallini and Zecchin correlation [5]

This correlation was obtained considering experimental data for R-22 and R134a and is the result of the study of the condensation of saturated vapours inside tubes.

$$h_{tp} = \left(\frac{k_f}{D_h}\right) * 0.05 * (Pr_l)^{0.333} * (Re_{eq})^{0.8} \quad \dots (1)$$

Where the equivalent Reynolds number, Reeq, is defined by

$$Re_{eq} = Re_g * \left(\frac{\mu_g}{\mu_l}\right) * \left(\frac{\rho_l}{\rho_g}\right)^{0.5} + Re_l \quad \dots (2)$$

1.2. Shah correlation [5]

This was developed using experimental data from the condensation of R-22 inside horizontal, vertical, and inclined pipes of different diameters. The correlation proposed defines the heat transfer coefficient as

$$h_{tp} = h_l * \left[ (1-X)^{0.8} + \left(\frac{3.8}{Pr^{0.38}}\right) * X^{0.76} * (1-X)^{0.04} \right] \quad \dots (3)$$

Where the liquid heat transfer coefficient is calculated by using the Dittuse-Boelter equation

$$h_l = 0.023 * \left(\frac{k_f}{D_f}\right) * (Pr_l)^{0.4} * (Re_l)^{0.8} \quad \dots (4)$$

1.3. Dobson and Chato correlation [5]

Dobson and Chato derived their correlation using R-22 and R-134a. They observed that the factors which control the flow in the case of condensation inside smooth horizontal tubes are gravity and vapour shear. At low vapour velocities, the effect of gravity prevails and condensate is formed in the upper part of the tube and flows downwards into a liquid pool which

advances throughout the tube (which is driven axially) due to vapour push (or flow) and to gravitational forces (or head).

For  $Re \leq 1250$ ,

$$Fr_{so} = 0.025 * (Re_l)^{1.59} * \left[ 1 + \left(\frac{1.09 * X_n^{0.034}}{X_n}\right) \right]^{1.5} * \left(\frac{1}{Ga}\right)^{0.5} \quad \dots (5)$$

And, for  $Re \geq 1250$ ,

$$Fr_{so} = 1.26 * (Re_l)^{1.04} * \left[ 1 + \left(\frac{1.09 * X_n^{0.034}}{X_n}\right) \right]^{1.5} * \left(\frac{1}{Ga}\right)^{0.5} \quad \dots (6)$$

The force-convective and the film-wise components are weighed so the first one is important at the bottom and the second at the top. Then the final expression is

$$Nu = \left(\frac{0.23 * Re_{vo}^{0.12}}{1 + 1.11 * X_n^{0.50}}\right) * \left(\frac{Ga * Pr_l}{Ja_l}\right)^{0.25} + \left[ 1 - \left(\frac{\theta}{\pi}\right) \right] * Nu_{forced} \quad \dots (7)$$

Where,  $\theta$  the angle subtended from the top of the tube to the liquid level,

$$Nu_{forced} = 0.0195 * (Re_l)^{0.8} * (Pr_l)^{0.4} * \phi(X_n) \quad \dots (8)$$

$$\phi(X_n) = \left[ 1.376 + \frac{C1}{X_n^{c2}} \right]^{0.5} \quad \dots (9)$$

Condition	C1	C2
$0 < Fr_f \leq 0.7$	$C1 = 4.172 + 5.48 * Fr_l$	7.242
$Fr_f > 0.7$	$C2 = 1.773 - 0.69 * Fr_l$	1.655

Under annular flow regime, Dobson and Chato derived a correlation by considering the Travis expression and assuming that the Reynolds number is usually greater than 1125. So the Nusselt number is given by a two-phase multiplier correlation:

$$Nu = 0.023 * (Re_l)^{0.8} * (Pr_l)^{0.4} * \left( 1 + \frac{2.22}{X_n^{0.89}} \right) \quad \dots (10)$$

2. Condensation in plate heat exchanger [2]

To facilitate the use of the plate heat exchanger as a Condense correlating equations for the dimensionless condensation heat transfer coefficient and friction factor based on the present data are provided. They are

$$Nu = \left(\frac{h_r * D_h}{\mu_l}\right) = 4.118 * (Re_{eq})^{0.4} * (Pr_l)^{0.333} \quad \dots (11)$$

And,

$$f_{tp} * Re^{0.4} * Bo^{-0.5} * \left(\frac{P_m}{P_c}\right)^{-0.8} = 94.75 * (Re_{eq})^{-0.00467} \quad \dots (12)$$

Where  $P_c$  is the critical pressure of R 134a,  $Re_{eq}$  is the equivalent Reynolds number and  $Bo$  is the boiling number defined as:

$$Re_{eq} = \left(\frac{G_{eq} * D_h}{\mu_l}\right) \quad \dots (13)$$

In which,

$$G_{eq} = G \left[ 1 - X_m + X_m \left(\frac{\rho_v}{\rho_l}\right)^{0.5} \right] \quad \dots (14)$$

**3. Condensation in flat channel with or without finned**

**3.1. Yang and Webb correlation [5]**

$$h = h_u \frac{A_u}{A} + h_f \frac{A_f}{A} \quad \dots (15)$$

Where,

$$h_f = 0.0265 * \left( \frac{k_f}{D_f} \right) * (\text{Pr}_i)^{0.4} * (\text{Re}_i)^{0.8} \quad \dots (16)$$

$$h_u = [h_{sh}^2 + h_{sf}^2]^{0.5} \quad \dots (17)$$

$$h_{sh} = 0.0265 * \left( \frac{k_f}{D_f} \right) * (\text{Pr}_i)^{0.4} * (\text{Re}_i)^{0.8} \quad \dots (18)$$

And,

$$h_{sf} = \frac{\tau_2 - \tau_1}{\frac{dp}{dz}} * k_f * \left[ \frac{1}{r_b} - \frac{1}{r_o} \right] * \text{Re}_{eq} * \frac{\text{Pr}_f^{0.333}}{W_e} \quad \dots (19)$$

**4. Condensation inside inclined Smooth tube [3]**

The general expression for the local condensing heat transfer is

$$h = \frac{(h_f * r_i * \theta) + (2\pi - \theta) * r_i * h_c}{2\pi * r_i} \quad \dots (20)$$

Where, hf and hc are condensation heat transfer coefficients for falling film in upper part and for axial flow in lower part, respectively.  $\theta$  is the falling film angle. They are calculated by the following equations.

$$h_f = 0.728 * \left[ \frac{(\rho_f - \rho_g) * g * h_{fg} * k_f^3}{\mu_f * d * (T_{sat} - T_w)} \right]^{0.25} \quad \dots (21)$$

$$h_c = 0.003 * \left( \frac{k_f}{\delta} \right) * f_i * (\text{Pr}_f)^{0.5} * (\text{Re}_f)^{0.74} \quad \dots (22)$$

$$\text{Re}_f = \left( \frac{4 * G * (1 - X) * \delta}{(1 - \varepsilon) * \mu_f} \right) \quad \dots (23)$$

And

$$f_i = 1 + \left( \frac{\mu_g}{\mu_f} \right)^{0.5} * \left[ \left( \frac{\rho_f - \rho_g}{\sigma} \right) * g * \delta^2 \right]^{0.25} \quad \dots (24)$$

The falling film angle  $\theta$  is obtained for each flow pattern as follows and where  $\theta_{start}$  is obtained by solving the following equation.

For annular, intermittent, and mist flow:

$$\theta = 0$$

For fully stratified flow:

$$\theta = \theta_{start}$$

$$A_L = \frac{d^2}{\delta} * [(2\pi - \theta_{start}) - \sin(2\pi - \theta_{start})] \quad \dots (25)$$

**5. Condensation in enhanced tubes**

**5.1. Cavallini et al. correlation [1] [3]**

This correlation is applicable to tubes which have a low fin, micro fin or grooved tubes geometry.

$$Nu = 0.05 * (\text{Re}_{eq})^{0.8} * (\text{Pr}_i)^{0.333} * R_x^S * (B_o * F_r)^Y \quad \dots (26)$$

Where,

$$\text{Re}_{eq} = 4 * M \left[ \frac{(1 - X) + X \left( \frac{\rho_l}{\rho_g} \right)^{0.5}}{\pi * d * \mu_l} \right] \quad \dots (27)$$

$$R_x = 2 * h * n_g * \left[ \frac{1 - \sin\left(\frac{\gamma}{2}\right)}{\pi * d * \cos\left(\frac{\gamma}{2}\right) + 1} \right] \quad \dots (28)$$

$$F_r = \left[ \frac{u^2 * G_o}{g * d} \right] \quad \dots (29)$$

**5.2. Yu and Koyama correlation [1]**

This correlation is used only for micro-fin tubes.

$$Nu = [Nu_f^2 + Nu_b^2]^{0.5} \quad \dots (30)$$

Where,

$$Nu_f = 0.152 * \left( \frac{\phi G}{X_w} \right) * \text{Re}_l^{0.68} [0.3 + 0.1 * \text{Pr}_i^{1.1}] \quad \dots (31)$$

$$\phi G = 1.1 + 1.3 \left[ \frac{G^{0.35} * X_w^{0.35}}{g * d * M * \rho_g (\rho_l - \rho_g)^{0.175}} \right] \quad \dots (32)$$

$$Nu_b = 0.725 * H(\varepsilon) * \left[ \frac{Ga * \text{Pr}_i}{\rho h_l * \eta_A} \right]^{0.25} \quad \dots (33)$$

$$H(\varepsilon) = \varepsilon + [10(1 - \varepsilon)^{0.1} - 8.0] * \varepsilon^{0.5} * (1 - \varepsilon^{0.5}) \quad \dots (34)$$

And

$$Ga = \left( \frac{g * \rho_l^2 * dM^3}{\mu_l^2} \right) \quad \dots (35)$$

**5.3. Condensation for annular and wavy flow regimes [4]**

Dobson et al. presented condensation heat transfer correlations for annular and wavy flow regimes. These correlations will be useful to compare refrigerants R-22 and R-134a. In case of R-134a 10% deviation found with experimental data and for R-22 deviation 6%.

For annular flow regimes correlation is:

$$Nu_{annular} = \frac{h.D}{k_l} = 0.06 \text{Re}_l^{0.8} \text{Pr}_l^{0.8} \frac{1}{X_w^{0.805}} \quad \dots (36)$$

$$h_{annular} \propto \frac{G^{0.8} k_l^{0.7} C_{pl}^{0.3}}{D^{0.2} \mu_l^{0.5}} \left( \frac{\rho_l}{\rho_v} \right)^{0.4} \left( \frac{\mu_v}{\mu_l} \right)^{0.08} R(x) \quad \dots (37)$$

where,

$$R(x) = x^{0.8} \left( \frac{1 - x}{x} \right)^{0.08} \quad \dots (38)$$

$$X'' = \left[ \frac{1-X}{X} \right]^{0.9} * \left( \frac{\rho_v}{\rho_l} \right)^{0.5} * \left( \frac{\mu_l}{\mu_v} \right)^{0.1} \quad \dots (39)$$

For wavy flow regimes correlation is:

$$Nu_{wavy} = \frac{h.D}{k_l} = 0.375 \left[ \frac{\rho_l(\rho_l - \rho_v)D_i^3 i_{lv}}{\mu_l \Delta T k_l} \right]^{0.25} \frac{1}{X''^{0.805}} \quad \dots (40)$$

$$h_{annular} \alpha \left[ \frac{\rho_l(\rho_l - \rho_v)D_i^3 i_{lv}}{\mu_l \Delta T k_l} \right]^{0.25} \left( \frac{\rho_l}{\rho_v} \right)^{0.115} \left( \frac{\mu_v}{\mu_l} \right)^{0.023} Q(x) \quad \dots (41)$$

where,

$$R(x) = \left( \frac{x}{1-x} \right)^{0.207} \quad \dots (42)$$

## II. CONCLUSION

This review has considered heat transfer investigations during in-tube and channel condensation. Possible research subjects have been summarized on the case in the literature, such as condensation heat transfer studies according to the tube orientation (horizontal, vertical, inclined tubes) and tube geometry (smooth and enhanced tubes), flow pattern studies of condensation.

In-tube condensation of refrigerants is a crucial event in many applications, such as condensers used in air conditioning, refrigeration, and heat pumps. Struggles with depletion of the stratospheric ozone layer and global warming caused by chemical compounds commenced relatively recently. To this end, condenser producers have been trying to change working fluids and to use enhanced geometries in the process of considering energy efficiency. These kinds of improvements will continue to take attention and will have an important impact on the HVAC industry in the future.

## REFERENCES

- [1] A.Cavallini, D. Del Col, L. Doretti, G.A. Longo and L. Rossetto, Heat transfer and pressure drop during condensation of refrigerants inside horizontal enhanced tubes, *International Journal of Refrigeration* 23 (2000), pp. 4-25.
- [2] Akio Miyara, Condensation of hydrocarbons – A review, *International Journal of Refrigeration* 31(2008), pp.621-632
- [3] J.R. Garcia-Cascales, F. Vera-Garcia, J.M. Corberan-Salvador & David Fuentes, Assessment of condensation heat transfer correlations in the modelling of fin and tube heat exchangers, *International Journal of Refrigeration* 30 (2007), pp. 1018-1028.
- [4] M. K.Dobson, J.C.Chato, S.P.Wang, D.K.Hinde & J.A.Gaibel, Initial Condensation Comparison of R 22 with R 134a & R-31. *Air Conditioning & Refrigeration Centre*, pp. 8-10.
- [5] Nae-Hyun Kim, Jin-Pyo Cho, Jung-Oh Kim and Baek Youn, Condensation heat transfer of R-22 and R-410A in flat aluminum multi-channel tubes with or without micro-fins, *International Journal of Refrigeration* 26(2003), pp. 830–839.



# Review on Available Studies Related to the Problem of Composite Plate with a Hole Subjected to Bending Loading

A. Nirav P Patel, B. Dharmendra S Sharma

**Abstract—** Composite materials become the attractive materials for the industries because of its properties and various applications. Holes and cutouts are bound to be present in any engineering structure because of its application, which cause serious problems of stress concentrations. This hole/opening works as stress raisers and may lead to the failure of the structure/machine component. The structure generally subjected to various types of loading like in-plane, bending and twisting. The review on available studies related to the problem of plate with various shaped holes subjected to bending loading is presented in this paper. The literatures related to complex variable approach, applied to the plate with a hole subjected to bending, are studied in this paper. The summary of method for solution of plate with a hole type of problems is concluded in this paper.

**Index Terms—:** Bending loading, complex variable approach, composites, plate with a hole.

## I. INTRODUCTION

In the modern era of technology, for aircrafts, space vehicles, under water transportation and many other lightweight applications, composite laminates are widely used. The strength and stiffness properties of composites can be controlled to meet the design requirements and this is the main feature of composites. Some benefits of composites materials are high strength, low maintenance requirement, resistance to the environment effects, and ability to be formed into a complex shape.

Nowadays, new types of composites like carbon fibers with higher strength and ultimate strain, thermoplastics matrices with its toughness, polymer matrix, metal matrix, and ceramic-matrix and carbon/carbon composites have been introduced and are widely used. Graphite/Epoxy,

A. Assistant professor at Mechanical Engineering Department, Gandhinagar Institute of technology, Khatraj-Kalol road, Village Moti Bhojan, Tal. Kalol, Gandhinagar-382721, India

Ph. D. student at Institute of Technology, Nirma University, Ahmedabad, Gujarat, India. Ph: +919824185450, Email: nirav\_npp@yahoo.com

B. Professor, Mechanical Engineering Department, Institute of technology, Nirma University, Ahmedabad, Ahmedabad, Gujarat, India. Email: dss\_iit@yahoo.com

Carbon/Epoxy, Boron/Epoxy, woven fabric are some of the examples of the composites which are also used in the industries. [1]

A lamina is the basic building block in fiber reinforced composite laminates. A lamina or ply is a plane or curved layer of unidirectional fibers or woven fabric in a matrix. It is also referred to as unidirectional lamina, which is an orthotropic material in its principal directions. A lamina is made up of two or more unidirectional lamina or plies stacked together at various orientations. The configuration indicating, in addition to the ply composition the exact location or sequence of various plies is called the stacking sequence. A symmetric laminate will have an identical lamina at a particular distance on either side of the mid plane. These laminates find wide applications because of their simplicity in construction and ease of analysis of the laminate.

## II. CHALLENGES TO DESIGN A COMPOSITE STRUCTURE

An orthotropic body has three mutual perpendicular plane of symmetry and anisotropic body has no plane symmetry. In the conventional and orthotropic material application of normal stress along principal material direction results in extension in the direction of applied stress and contraction in the perpendicular direction to the applied stress. Shear stress causes shearing deformation. For isotropic materials, the shear modulus is independent of other material properties. For anisotropic materials, the application of normal stress leads not only to extension in the direction of the stress and contraction perpendicular to it, but also to shearing deformation. Same way, shearing stress causes extension and contraction in addition to the distortion of shearing deformation. The effect of shear- extensional coupling is the very important characteristic of composites. The static analysis of symmetric laminates uncouples into two different problems such as pure in-plane deformation with pure in-plane forces and the other, out of plane deformation with bending and transversal forces. However, in case of unsymmetric laminates, the bending extension-coupling will occur because of which, the shear and normal forces acting on plate result in not only the in-plane deformations, but also twisting and bending producing plate curvatures and vice-versa. This type of coupling behavior is the key to design the composites. It is a challenge for designers to control this

behavior when the structure placed into the practical application.

Holes are generally made into the structure because of its practical application like boiler manhole, submarine applications, boats etc. These holes and cutouts act as a stress raisers and because of which the material fails. This is the critical problem for the designers. Extensive studies have been made on stress analysis of isotropic materials globally and find applications. The literatures related to the stress analysis of plate with a hole subjected to the bending are studied and reviewed in this paper.

### III. ANALYTICAL METHODS FOR THE STRESS ANALYSIS OF COMPOSITES PLATE SUBJECTED TO BENDING

Many researchers have tried for stress analysis around holes in composite plates. A good content of studies related to stress analysis of isotropic material, the plate with a hole subjected to in-plane loading is available but a very few literature is available in the study of stress analysis of a plate with a hole subjected to bending loading.

Goodier [2] has given the solution for the plate with circular and elliptical hole subjected to bending for the isotropic material. Reissner [3] has also given the solution for the isotropic plate subjected to bending.

Lekhnitskii [4] and Savin [5] has given the generalized solution of plate with hole subjected to bending loading. They both have considered the various shaped holes like square, triangle, elliptical and circular.

In 1862 Airy [6] has given that the stress might be expressed by means of one single auxiliary function. This function is called as Airy function. According to Airy [6] the stress may be expressed in the following manner in “(1)”:

$$\begin{aligned}
 X_x &= \frac{\partial^2 U}{\partial y^2}; \\
 X_y &= -\frac{\partial^2 U}{\partial x \partial y}; \\
 Y_y &= \frac{\partial^2 U}{\partial x^2}
 \end{aligned}
 \tag{1}$$

U(x, y)=Airy function  
 X<sub>x</sub>=Stress in X direction  
 X<sub>y</sub>=Shear stress  
 Y<sub>y</sub>= Stress in X direction

J. C. Maxwell was the first person to notice that the Airy function must satisfy the “(2)” given below:

$$\frac{\partial^4 U}{\partial x^4} + 2\frac{\partial^4 U}{\partial x^2 \partial y^2} + \frac{\partial^4 U}{\partial y^4} = 0
 \tag{2}$$

Equation (2) is called biharmonic and its solutions are called as biharmonic functions.

Mushkelishvili's [7] has given that every biharmonic function U(x, y) of the two variables x, y might be represented by the functions of the complex variable z=x+iy.

This fact is very important because, the complex variable functions are biharmonic functions and it satisfies the “(2)”.

Mushkelishvili's [7] has given the Airy function U as shown below (refer “(3)”):

$$2U = \bar{z}\phi(z) + z\overline{\phi(z)} + \chi(z) + \overline{\chi(z)}
 \tag{3}$$

$\phi, \chi$  are “arbitrary” functions.

Above “(3)” was also obtained by E. Goursat [8] but the form of the function is little bit different.

Lekhnitskii [4] and Savin [5] both have used the Mushkelishvili's [7] complex variable approach to find out the solution for the plate with hole subjected to bending loading. .

Savin [5] has given the results mainly of the isotropic materials. Lekhnitskii [4] has given the results for the isotropic as well as anisotropic plates. Lekhnitskii [4] has given the results for the plywood plates. Results for a circular hole in plywood plate were given for bending, all round bending and twisting in his work. The results for moment distribution around a circular hole in ply wood and isotropic steel plate given by Lekhnitskii [4] are given in Fig. 1.

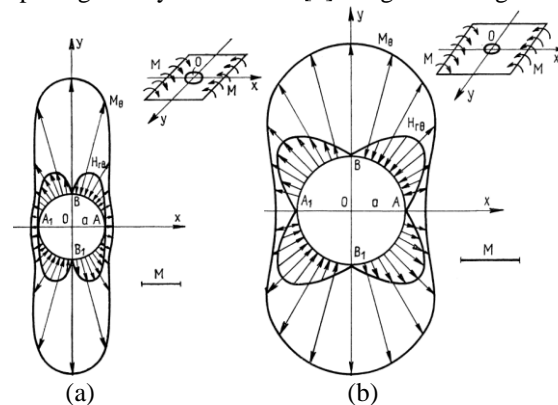


Fig.1 (a) Stress distribution in plywood plate [4]  
 (b) Stress distribution in steel plate [4]

Lekhnitskii's [4] majority of the work was concentrated in the isotropic and orthotropic plates. Symmetric laminates and unsymmetric laminates were not addressed in Lekhnitskii's [4] work.

Based on Mushkelishvili's [7] complex variable approach Ukadgaonker and Rao [9] has given the generalized solution for composite plate with arbitrary shaped hole subjected to bending. Ukadgaonker and Rao [9] has studied the moment distribution around circular, elliptical, triangular, square, rectangular and several irregular holes in various cross ply and angle ply symmetric laminates. The study revealed that the moment around the hole will decrease for the laminates with the large number of ply groups and vice versa. The results also revealed that the moment and symmetry of the

distribution depends on the combined effect of all the parameters, like hole geometry, loading, young’s moduli and Poisson’s ratios of the material and the flexural moduli of the laminates.

Ukadgaonker and Rao [9] has adopted Gao’s [10] condition to reduce the superposition of the solutions of the uniaxial loading. These conditions consider any arbitrary orientation of uniaxial and biaxial moments. This was achieved by introducing the biaxial loading factor  $\lambda$  and orientation angle  $\beta$  into the boundary condition at infinity. The plate with a hole subjected to bending is as shown in Fig. 2. The boundary conditions were obtained as in “(4)”:

$$\begin{aligned}
 M_x^\infty &= \frac{M}{2} [(\lambda + 1) + (\lambda - 1) \cos 2\beta] \\
 M_y^\infty &= \frac{M}{2} [(\lambda + 1) - (\lambda - 1) \cos 2\beta] \\
 M_{xy}^\infty &= \frac{M}{2} [(\lambda - 1) \sin 2\beta]
 \end{aligned}
 \tag{4}$$

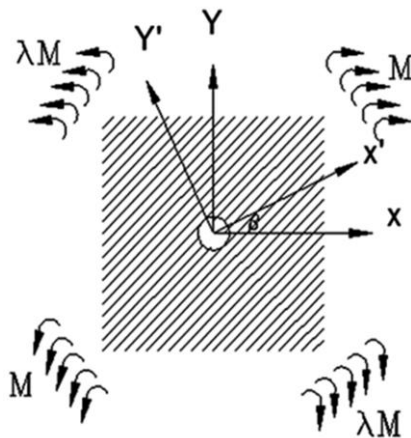


Fig. 2 Moment around hole

Ukadgaonker and Rao [9] have distributed the problem of plate with a hole subjected to bending into two stages. For each stage the stress functions were obtained. The stages are as shown in Fig. 3.

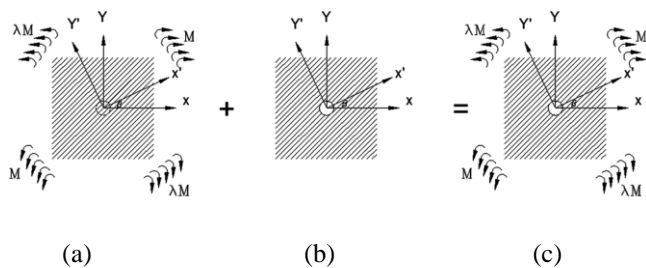


Fig. 3 The scheme of solution [9]

- (a) Remote loading on hole free plate: Stage 1 [9]
- (b) Plate with hole: Stage 2 [9]
- (c) Remote loading on plate with hole [9]

In the first stage the stress functions were obtained for the hole free plate due to applied moment at infinity. The boundary conditions from these stress function were obtained on the fictitious hole. For the second stage the plate with hole was applied by a negative of the boundary condition obtained from the first stage. The second stage stress functions were obtained from these boundary conditions. By adding these two stage stress functions the last stress functions were obtained. These obtained stress function were put into the generalized moment equations shown below [9]:

$$M_x = -2 \operatorname{Re} [p_1 \phi'(z_1) + q_1 \psi'(z_2)] \tag{5}$$

$$M_y = -2 \operatorname{Re} [p_2 \phi'(z_1) + q_2 \psi'(z_2)] \tag{6}$$

$$M_{xy} = -2 \operatorname{Re} [p_3 \phi'(z_1) + q_3 \psi'(z_2)] \tag{7}$$

Where,

$$p_1 = D_{xx} + D_{xy} s_1^2 + 2D_{xs} s_1;$$

$$p_2 = D_{xy} + D_{yy} s_1^2 + 2D_{ys} s_1;$$

$$p_3 = D_{xs} + D_{ys} s_1^2 + 2D_{ss} s_1$$

$$q_1 = D_{xx} + D_{xy} s_2^2 + 2D_{xs} s_2;$$

$$q_2 = D_{xy} + D_{yy} s_2^2 + 2D_{ys} s_2;$$

$$q_3 = D_{xs} + D_{ys} s_2^2 + 2D_{ss} s_2$$

$s_1$  and  $s_2$  are constants of anisotropy obtained as roots of characteristic equation shown below [10]:

$$D_{yy} s^4 + 4D_{ys} s^3 + 2(D_{xy} + 2D_{ss}) s^2 + 4D_{xs} s + D_{xx} = 0 \tag{8}$$

The flexural moduli were obtained based on Tsai and Hahn [11].

$$D_{ij} = \frac{8}{n^3} \sum_{t=1}^{n/2} Q_{ij}^{(t)} [t^3 - (t-1)^3] \tag{9}$$

Where,  $n$  = total number of plies in the laminate

$t$  = weighting factor

$Q_{ij}$  ( $i, j = x, y, s$ ) = Stiffness coefficients of a given ply group

Ukadgaonker and Rao [9] have used the concept of conformal mapping to transfer the complicated shapes to simpler one. By using affine transformation, the area external to a given region in  $z$  plane is transform by the area outside the unit circle in  $\zeta$  plane using the given equation [9]:

$$z = \omega(\zeta) = R\left(\zeta + \sum_{k=1}^n \frac{m_k}{\zeta^k}\right) \tag{10}$$

Ukadgaonker and Rao [9] have studied the effect of stacking sequence. Results were obtained for the 4, 8 and 16 ply groups containing various shaped hole subjected to different types of loading condition. Some of the results are shown in Fig. 4 and Fig. 5. Fig.4 represents the moment distribution around the circular hole for different stacking sequence. Fig.5 represents the moment distribution around a square hole.

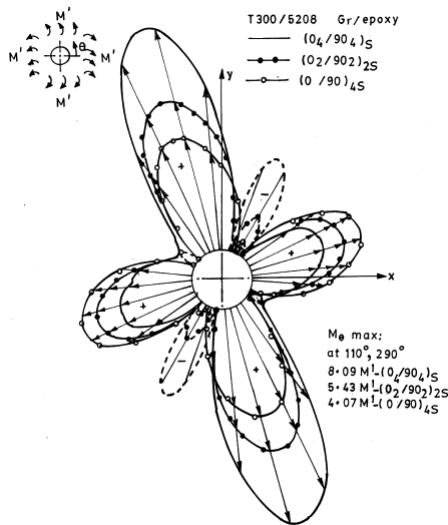


Fig.4 Moment around circular hole [9]

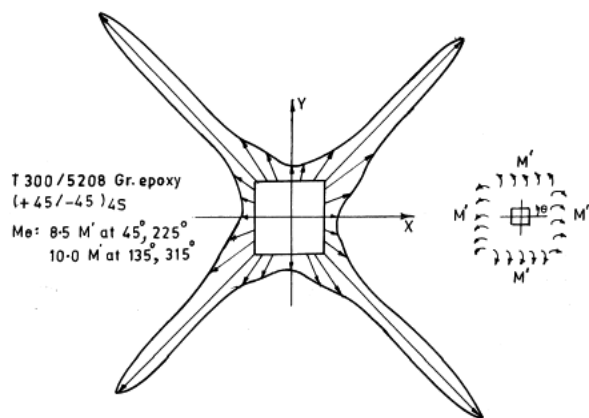


Fig.5 Moment around square hole [9]

Sharma and Patel [12] have used Mushkelishvili's [7] complex variable approach and have given the generalized solution for the moment distribution around the different hole. Gao's [10] condition was used to reduce the superposition of the solutions of the uniaxial loading. The concept of conformal mapping to transfer the complicated regions to a simpler circular surface was used. Schwartz integral was

applied to find the boundary conditions around the circular hole.

Sharma and Patel [12] have obtained the moment around the circular, elliptical and triangular hole in various cross ply and angle ply symmetric laminates. The effect of various parameters like loading factor, stacking sequence, fiber orientation, flexural moduli and hole geometry, was studied on the 16 -ply cross ply and 16 -ply angle ply laminates of graphite/epoxy, glass/epoxy, boron/epoxy etc. with circular, elliptical and triangular shaped hole, for the cylindrical bending, equi-biaxial bending and twisting moments. The results were evident that the maximum value of normalized tangential moment decreases as number of ply groups increases. The stacking sequence and the loading condition have significant effect on the distribution of moment around cutouts. The maximum value of moment occurred for the boron/epoxy material than the graphite/epoxy and glass/epoxy. The absolute value of normalized tangential moment is increase as the corner radius of triangular hole decreased.

The problems of anisotropic plates containing holes, cracks and/or inclusions have been studied extensively for two dimensional deformations by Hsieh and Hwu [13, 14, 15]. Stroh-like formalism for the bending theory of anisotropic plates was given. By this newly developed formalism, most of the relations for bending problems can be organized into the forms for two-dimensional problems. The forces and the moment around the hole boundary is given by the following “(11)” and the one of the results obtained is shown in Fig. 6 [14]:

$$N_n = N_{ns} = M_n = 0$$

$$N_s = \cos \theta \left\{ \sigma_1^{(0)} - c \sigma_1^{(3)} + \sigma_2^{(1)} \right\} + \sin \theta \left\{ \sigma_2^{(0)} - \frac{1}{c} \sigma_2^{(3)} + \sigma_1^{(1)} \right\}$$

$$M_s = \cos \theta \left\{ \sigma_1^{(0)} - c \sigma_1^{(3)} + \sigma_2^{(1)} \right\} + \sin \theta \left\{ \sigma_2^{(0)} - \frac{1}{c} \sigma_2^{(3)} + \sigma_1^{(1)} \right\};$$

$$M_{ns} = \cos \theta \left\{ \sigma_1^{-(0)} - c \sigma_1^{-(3)} + \sigma_2^{-(1)} \right\} + \sin \theta \left\{ \sigma_2^{-(0)} - \frac{1}{c} \sigma_2^{-(3)} + \sigma_1^{-(1)} \right\}; \tag{11}$$

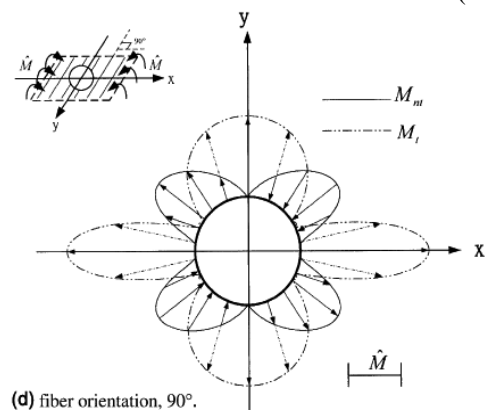


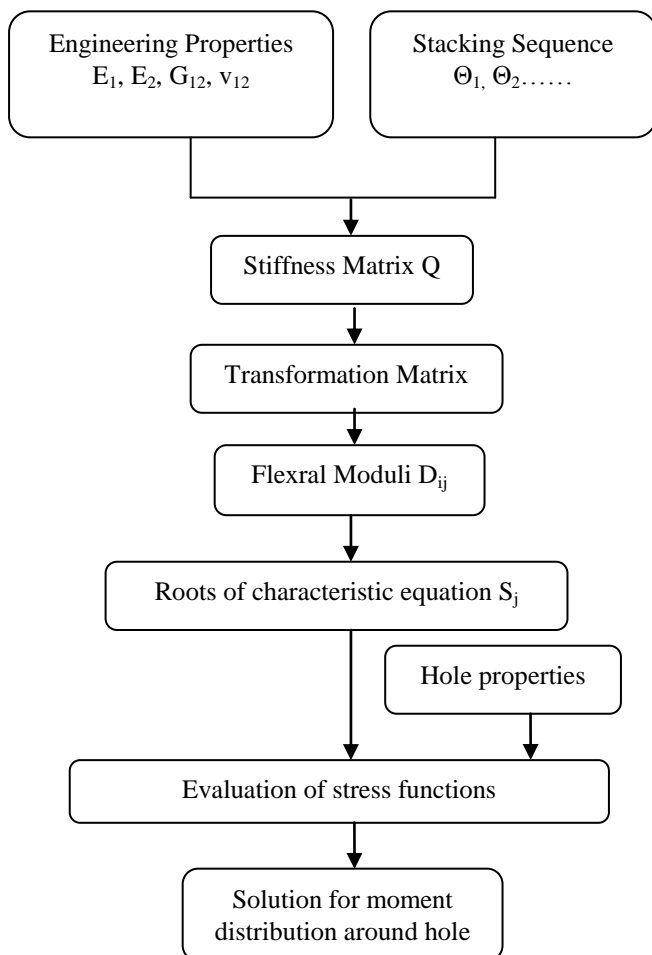
Fig. 6 Moment distribution around a circular hole for anisotropic plate [13]

An analytical investigation of the effects of holes on the moment distribution of symmetric composite laminates

subjected to bending moments was described by Prasad and Shuart [16]. Laminates with a circular hole or with an elliptical hole were studied subjected bending and twisting.

IV. CONCLUSION:

Nowadays the composite materials become the best alternative against the conventional materials because of its features. It's big challenge for the engineers to design these materials, especially for the problems of plate with a hole. These problems can be solved by different methods like FEA, FEM, complex variable method. The complex variable method given by Muskhelishvili is the best method to solve this type of problems because of its speed for solution, no messing requirement, gives solution at every point around a hole, easy to handle etc. The summary of this complex variable method is as shown below:



REFERENCES

[1] M. Denial Isaac and Ori Ishai, *Engineering mechanics of composite materials*. Oxford University press, USA, 2006.  
 [2] J. N. Goodier, "The influence of circular and elliptical holes on the transverse flexure of elastic plates," *Phil Mag*, Ser 7 1936, pp. 22:69-80.

[3] E. Reissner, "The effect of transverse shear deformation on the bending of elastic plates," *J Appl Mech, Trans ASME*, 1945, pp. 67:A69-77.  
 [4] S. G. Lekhnitskii, *Anisotropic plates*. New York, Gordon and Breach, 1968.  
 [5] G. N. Savin, *Stress concentration around holes*. New York Pergamon Press, 1961.  
 [6] G. B. Airy, "On the strains in the interior of Beams," *Philosophical Transactions of the Royal Society*, pp. 153:49–80, 1863.  
 [7] N. I Muskhelishvili, *Some basic problems of the mathematical theory of elasticity*. P. Noordhoff Ltd., Groningen, the Netherlands, 1963.  
 [8] E. Goursat, "Sur l'equation  $\Delta\Delta u = 0$ ," *Bull. De la Soc. Math. Soc.*, vol. 26, pp. 236, 1898.  
 [9] V.G. Ukadgaonker, and D.K.N Rao, "A general solution for moments around holes in symmetric laminates," *Composite Structures*, Vol. 49, pp. 41-54, 2000.  
 [10] X. L. Gao, "On the complex variable displacement method in plane isotropic elasticity", *Mechanics Research Communications*, Vol. 31, pp. 169–173, 2004.  
 [11] S. W. Tsai, H. T. Hahn, *Introduction to Composite Materials*, Technomic Publishing, Lancaster, 1980.  
 [12] D. S. Sharma and N. P. Patel, "Moment distribution around circular/elliptical/triangular cutouts in infinite symmetric laminated plate," *Fourth International Conference on Structural Stability and Dynamics (ICSSD)*, pp. 818-824, 4–6 January, 2012.  
 [13] M.C. Hsieh and C. Hwu, "Anisotropic elastic plates with holes/cracks/inclusions subjected to out-of-plane bending moments," *International Journal of Solids and Structures*, vol. 39 (19), pp. 4905–4925, 2002a.  
 [14] M.C. Hsieh and C. Hwu, "Explicit expressions for the fundamental elasticity matrices of Stroh-like formalism for symmetric/unsymmetric laminates," *The Chinese Journal of Mechanics, Series A* vol. 18 (3), pp. 109–118, 2002b.  
 [15] C. Hwu, "Stroh-Like complex variable formalism for bending theory of anisotropic plates," *International Journal of Solids and Structures* vol. 40, pp. 3681–3705, 2003.  
 [16] C. B. Prasad and M. J. Shuart, "Moment Distributions Around Holes in Symmetric Composite Laminates Subjected to Bending Moments," *aiaa journal*, vol. 28, no. 5, may 1990.

# Improving the Performance of Fodder Cutter by using Three Blades Instead of Two Blades

A. Amruta N. Mahajan, B. D.B. Sadaphale, C. Sapna A. Solanki

**Abstract**—Chaff cutter, a mechanical device for cutting straw or hay into small pieces, before being mixed together with other forage and fed to horses and cattle. This aids the animal's digestion and prevents animals from rejecting any part of their food. A detailed study of injuries and machine characteristics resulted in a safer fodder-cutter design. The design changes are cost effective and can be incorporated, in both existing and new fodder-cutter machines. Chaff cutters have evolved from the basic machines into commercial standard machines that can be driven at various speeds and can achieved various lengths of cuts of chaff with respect to animal preference type. The improvement in design is made by considering various parameters such as cost, efficiency, quality, life, maintenance of it, etc. The work is going on to find out the amount of power/Energy loss in transmission & also Energy losses of cutting wheel during each shearing, Minimum cutting velocity required for clean shearing, Optimum blade angle (Blade geometry), Optimum cutting clearance between guide cutting edge and cutter. To get desired output the model of the blade is created using modeling and design software Solid Edge V20 and analysis is done using Ansys 11 to find out the energy losses during each cut & optimum blade angel.

**Index Terms**—Blade angle, Chaff cutter, Forage, stem properties

## I. INTRODUCTION

INDIA is an agriculture based Country and our 90% economy is depends upon Agriculture product. There are different crops produced throughout the year. After the crop cutting and the grain separate out the remaining stem or chaff is used for different application like cattle food, or making bio mass for general cooking water heating, and many more application. But it is majorly used as cattle food. For that it is cut by using chaff cutter and hence called as Fodder-cutter machines or stem cutter machine. These are used every day by farmers and their families in India for preparation of fodder for the livestock they own.

Actually, Chaff cutter is a mechanical device for cutting straw or hay into small pieces — before being mixed together

with other forage and fed to horses and cattle. This aids the animal's digestion and prevents animals from rejecting any part of their food.

Chaff and hay played a vital role in most agricultural production as it was used for feeding horses. Horses were extensively used in farming operations until they were replaced by tractors in the 1940s.

Chaff cutters have evolved from the basic machines into commercial standard machines that can be driven at various speeds and can achieved various lengths of cuts of chaff with respect to animal preference type. New chaff cutter machines include portable tractor driven chaff cutter - where chaff cutter can be in the field and load trolleys (if required).

In olden days, the chaff box or cutter was a simple device having *single blade* for cutting straw chaff, hay, and oats into small pieces. The chaff box was made largely of wood with only a small amount of iron work, it cost relatively little to make. But they were human operated.

At the end of nineteenth century modernize Chaff Cutters were made available in market which was more efficient and require less power for operation and they were with two blades. The main components of these fodder cutter are following: Flywheel, Cutting Blade, Motor with variable speed

Such stem cutters are used commonly as these are motorized & safe. We focus on improving the performance of such machines by replacing two cutting blades with three blades as shown in fig.1



Fig. 1. Three blade chaff cutter flywheel.

- M.E.(machine Design) Student of SSBT's College Of Engineering,Bambhori,Jalgaon-(e-mail: amruta8581@gmail.com).
- Works with the mechanical Engg Department of SSBT's College Of Engineering, Bambhori,Jalgaon(e-mail: devendra\_sadaphale@rediffmail.com).
- Asstworks with the mechanical Engg Department of Hasmukh Goswami College Of Engineering,Ahmedabad(e-mail: sapana49@gmail.com).

II. METHODOLOGY

The study was done in two phases. The following are the different points which are considered to improve the design of the stem cutter.

Firstly, the stem samples categorized according to the % moisture content i.e. High moisture stem, Medium moisture stem & low moisture stem. For these different types of stems , their properties such as density of stem, hardness, compressive strength, tensile strength, modulus of elasticity etc. are found out experimentally.

Secondly, solid model of blades & flywheel are created using modeling and design software Solid Edge V20 and analysis is done using Ansys 11 to find out the energy losses during each cut & optimum blade angle, vibration analysis, deformation during cutting etc.

III. EXPERIMENTAL CALCULATION FOR FINDING PROPERTIES OF STEMS

Sample stem selected are of 8cm in length. Samples chosen for analysis are of following three types

- 1) High Moisture Stem
- 2) Medium Moisture Stem
- 3) Low Moisture Stem

A. Moisture Content ( H.M. Stem)

--Initial Weight (I.W.) = 43.440 grams.

--Final Weight (F.W.) = 15.530 grams.

$$\text{Moisture Content in \%} = \frac{I.W - F.W}{(I.W.)} \times 100$$

$$= \frac{43.440 - 15.530}{43.440} \times 100$$

Moisture Content in % = 64.25 %

Similarly,

For Medium Moisture Stems = 41.74%

For Low Moisture Stems = 7 %

B. Density (H.M.stem)

-- Stem Sample Length = 8cm

-- Initial Weight with Wax, m = 26.120 grams

-- Rises Level of water, V = 25ml

$$\text{Density} = \frac{\text{Mass}}{\text{Volume}} = \frac{m}{V}$$

$$= \frac{26.120}{25}$$

=1.0448 grams / ml

=1.0448 X 10<sup>3</sup> Kg/ ml

Density =1044.8 X 10<sup>6</sup> Kg / m<sup>3</sup>

Similarly,

Density for Medium Moisture Stem= 1055 x 10<sup>6</sup> kg/m<sup>3</sup>

Density For Low Moisture Stems = 1133.7 X 10<sup>6</sup> Kg/m<sup>3</sup> can be calculated

C. Hardness

For finding hardness of all types of stems using Rockwell hardness testing m/c with load 60 kg

--At right to center = 46

--At Joint = 40



Fig. 2. Rockwell hardness testing on High moisture Stem Sample

D. Compressive Strength

For finding the compressive strength, the sample is tested on compression testing machine as shown in fig.2(for 120 kg load)

-- H.M. Stem = 1177.2 N

-- M.M. Stem = 569 N

-- L.M. Stem = 108 N



Fig. 3. Compression testing on High moisture Stem Sample

E. Young’s Modulus (Modulus of elasticity)

Tensile testing on the stem sample is done to find out Young’s Modulus

Tensile Strength (Tensile Force) (Apply Tensile Force Up To Break)  $m = 236 \text{ kg}$

For high moisture stem,

Original Diameter (O.D.) = 17mm

Original Length (O.L.) = 501mm

Change in Diameter (C.D.) = 13.5mm &

Change in Length (C.L.) = 505mm

$$r = 17/2 = 8.5\text{mm} = 8.5 \times 10^{-3} \text{ m}$$

$$\text{Tensile Stress} = F/A_o = mg / \pi r^2$$

$$\text{Tensile Stress} = 1.02 \times 10^7 \text{ N/m}^2$$

$$\text{Tensile Strain} = \frac{\text{Increases In Length}}{\text{Original Length}}$$

$$\text{Tensile Strain} = \frac{C.L. - O.L.}{O.L.} = \frac{505 - 501}{501}$$

$$\text{Tensile Strain} = 0.008$$

$$\text{Young’s Modulus} = \frac{\text{Tensile Stress}}{\text{Tensile Strain}}$$

$$= \frac{1.02 \times 10^7}{0.008}$$

$$\text{Young’s Modulus} = 1.275 \times 10^9 \text{ N/m}^2$$

Young’s modulus For M.M. stem =  $0.873 \times 10^9 \text{ N/m}^2$

Young’s modulus For L.M. stem =  $0.336 \times 10^9 \text{ N/m}^2$

F. Poisson’s Ratio

$$\text{Poisson’s Ratio} = \frac{\text{Lateral Strain}}{\text{Longitudinal Strain}}$$

$$= \frac{\text{Lateral Strain}}{\text{Tensile Strain}}$$

$$\text{Lateral Strain} = \frac{\text{Decreases In Dia.}}{\text{Original Dia.}}$$

$$= \frac{O.D. - C.D.}{O.D.}$$

$$= 0.206$$

Longitudinal Strain (Tensile Strain) = 0.008 (it is already calculated as above)

$$\text{Poisson’s Ratio} = \frac{\text{Lateral Strain}}{\text{Tensile Strain}}$$

$$= \frac{0.206}{0.008}$$

$$\text{Poisson’s Ratio} = 25.75$$

For High Moisture Stem = 25.75

For Medium Moisture Stem = 25.31

For Low Moisture Stem = 15.94

IV. ANALYSIS OF THE BLADE AND FLYWHEEL ASSEMBLY FOR THREE SPOKE AND TWO SPOKE

For the analysis we consider the three different blade materials as high carbon steel, medium carbon steel & low carbon steel. Before doing analysis, it is necessary to do meshing of the assembly as shown in fig below.

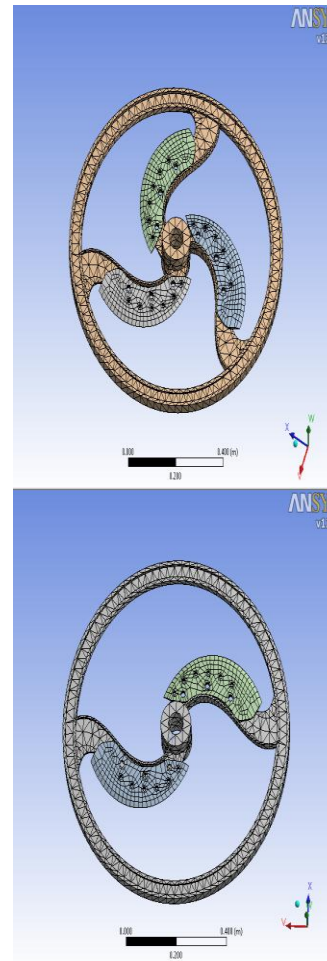


Fig. 4. Meshing of the 3 & 2 spike flywheel and blade



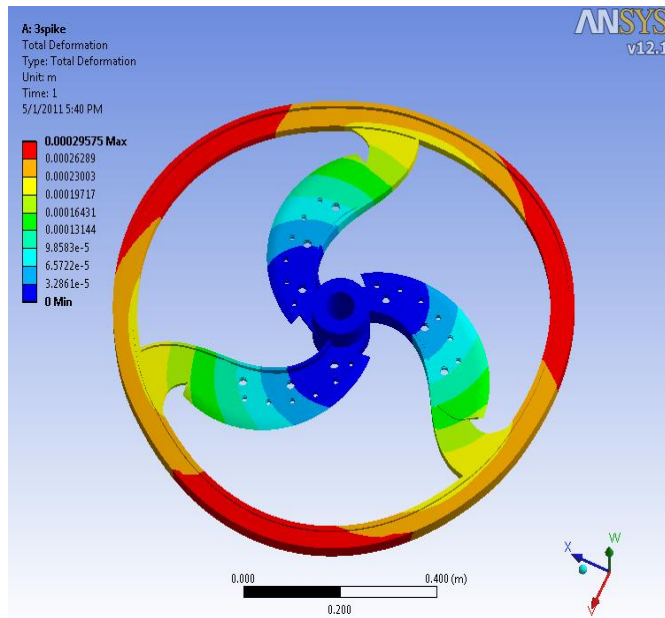


Fig.5 Deformation of 3 spike flywheel and blade

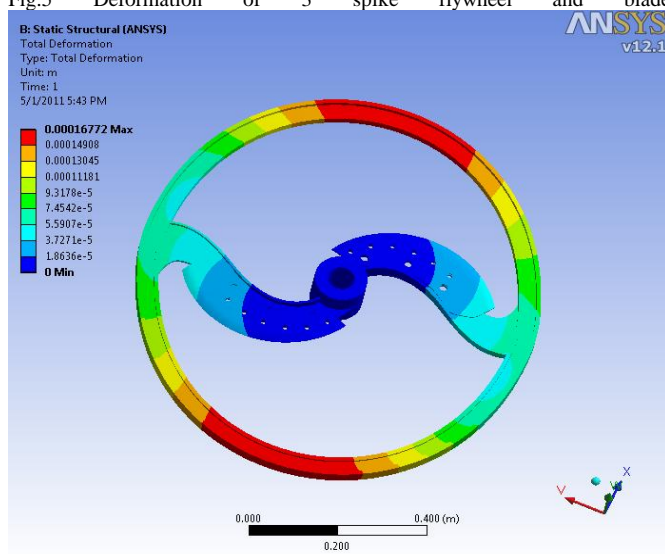


Fig. 6 Deformation of 2 spike flywheel and blade

Vibration Analysis also performed on different blade materials for different stem properties.

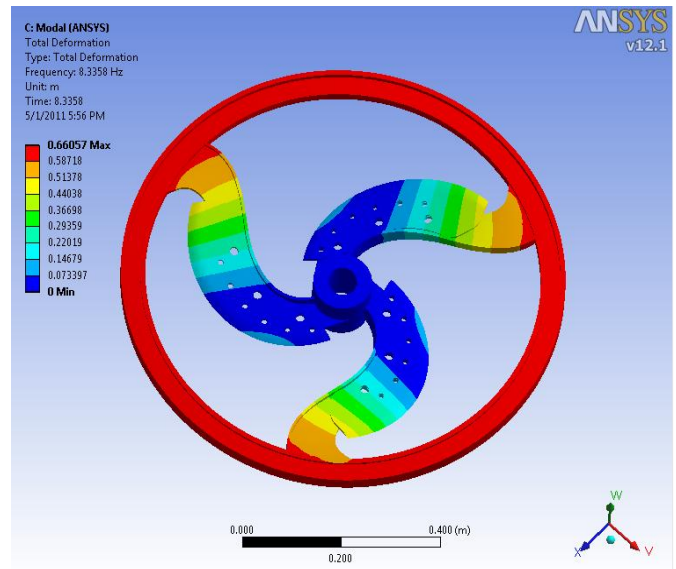


Fig. 7 vibration deformation for high carbon steel blade with 3 spoke

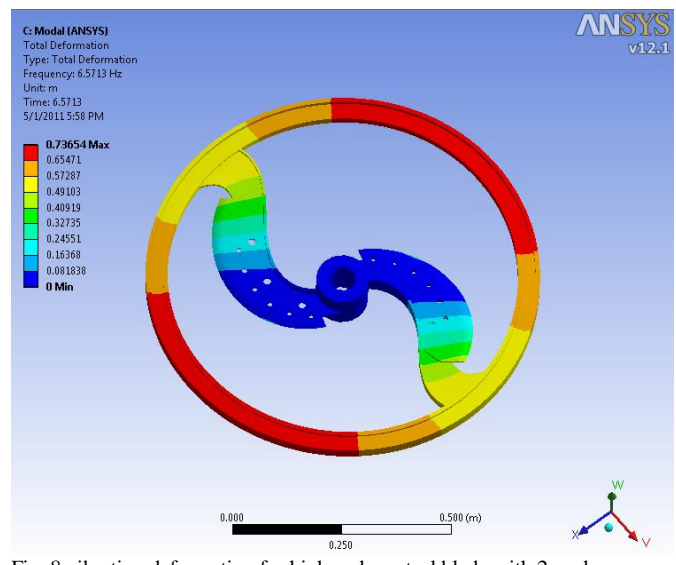


Fig. 8 vibration deformation for high carbon steel blade with 2 spoke

VI CONCLUSION

From the analysis of two & three blades cutter following results are obtained. So, we can say that three blade fodder cutter improves the performance as it is having less vibrational deflection though total Deflection and Von-mises stress are more for 3 spike for all material condition, because of more centrifugal force is acting on its components.

TABLE I  
 RESULT S FOR HIGH CARBON STEEL BLADE

parameter	2 Blade Cutter	3 Blade Cutter
Total deflection(m)	0.000167	0.00029
Von-mises stress(pa)	4.1e5	7.38e5
Max. Shear stress(pa)	2.12e5	3.8e5
Vibrational Deflection(m)	0.73	0.66

TABLE II  
RESULTS FOR LOW CARBON STEEL BLADE

parameter	2 Blade Cutter	3 Blade Cutter
Total deflection(m)	0.000637	0.00011218
Von-mises stress(pa)	4.096e5	7.488e5
Max. Shear stress(pa)	2.14e5	3.865e5
Vibrational Deflection(m)	1.84	1.65

TABLE III  
RESULTS FOR MEDIUM CARBON STEEL BLADE

parameter	2 Blade Cutter	3 Blade Cutter
Total deflection(m)	0.000244	0.00043
Von-mises stress(pa)	4.16e5	7.4e5
Max. Shear stress(pa)	2.13e5	3.85e5
Vibrational Deflection(m)	0.83	0.734

Due to large mass density the 3 spike is more stable from vibrational point of view as it is affected by less modal deflection because of higher weight.

Stress are in the following range

Low carbon steel > Medium carbon steel > High C steel

Stresses and Deflection will be least for High C steel because of higher yield strength.

#### REFERENCES

- [1] K.S.Zakiuddin," Formulation of data based ANN model for the Human powered fodder – chopper," Journal of Theoretical and Applied Information Technology, Islamabad pp.79-85,2010
- [2] Dinesh Mohan," Development of safer fodder-cutter machines: a case study from north India", Safety Science 42 ,pp.43–55,2004
- [3] Yiljep Y. P. and Mohammad U. S., "Effect of knife velocity on cutting energy and efficiency during impact cutting of sorghum stalk." *Agril. Engg. International: the CIGR EJournal*, 7: 1- 8,2005
- [4] Vivek P. Pathak, "Effect Of Different Cutters, Feed Rates And Speed On Performance Of Chopping Of Green Sorghum,".
- [5] T.K.Reza, "Paddy Stems Cutting Energy And Suggested Blade Optimum Parameters," Pakistan Journal of Biological Sciences 10(24),pp.4523-4526,2007.
- [6] Jekendra, Y., " Physical and rheological properties of forage Crops with reference to cutting" Arch. Zootec. 48:pp. 75-78,1999

# A Review on Preparation Methods & Thermal Conductivity of Nanofluids

A. Milind A. Patel, B. Ankit K. Patel

**Abstract**— Nanofluids are nanotechnology based heat transfer fluids that are derived by stably suspending nanometer-sized particles in conventional heat transfer fluids-usually liquid. This paper summarizes the study of nanofluids, such as the preparation methods of nanofluids, thermal conductivity & effect of some parameters on thermal conductivity of nanofluids. At last, the paper identifies the opportunities for future research.

**Index Terms**— Heat Transfer, Nanofluid, Thermal Conductivity.

## I. INTRODUCTION

There are several methods to improve the heat transfer efficiency. Some methods are utilization of extended surfaces, application of vibration to the heat transfer surfaces, and usage of micro channels. Heat transfer efficiency can also be improved by increasing the thermal conductivity of the working fluid. Commonly used heat transfer fluids such as water, ethylene glycol, and engine oil have relatively low thermal conductivities, when compared to the thermal conductivity of solids. High thermal conductivity of solids can be used to increase the thermal conductivity of a fluid by adding small solid particles to that fluid. The feasibility of the usage of such suspensions of solid particles with sizes on the order of millimeters or micrometers was previously investigated by several researchers and significant drawbacks were observed. These drawbacks are sedimentation of particles, clogging of channels and erosion in channel walls, which prevented the practical application of suspensions of solid particles in base fluids as advanced working fluids in heat transfer applications [1], [2]. Nanofluids are a new class of fluids engineered by dispersing nanometer-sized materials (nanoparticles, nanofibers, nanotubes, nanowires, nanorods, nanosheet, or droplets) in base fluids. In other words, nanofluids are nanoscale colloidal suspensions containing condensed nanomaterials. They are two-phase systems with one phase (solid phase) in another (liquid phase). Nanofluids

have been found to possess enhanced thermo physical properties such as thermal conductivity, thermal diffusivity, viscosity, and convective heat transfer coefficients compared to those of base fluids like oil or water. It has demonstrated great potential applications in many fields.

## II. PARTICLE MATERIAL & BASE FLUID

Many different particle materials are used for nanofluid preparation. Al<sub>2</sub>O<sub>3</sub>, CuO, TiO<sub>2</sub>, SiC, TiC, Ag, Au, Cu, and Fe nanoparticles are frequently used in nanofluid research. Carbon nanotubes are also utilized due to their extremely high thermal conductivity in the longitudinal (axial) direction. Nanoparticles used in nanofluid preparation usually have diameters below 100 nm. Spherical particles are mostly used in nanofluids. However, rod-shaped, tube-shaped and disk-shaped nanoparticles are also used.

Base fluids mostly used in the preparation of nanofluids are the common working fluids of heat transfer applications; such as, water, ethylene glycol and engine oil. In order to improve the stability of nanoparticles inside the base fluid, some additives are added to the mixture in small amounts.

## III. PRODUCTION METHODS

### A. Production of Nanoparticles

Production of nanoparticles can be divided into two main categories, namely, physical synthesis and chemical synthesis. Yu et al. [5] listed the common production techniques of nanofluids as follows.

**Physical Synthesis:** Mechanical grinding, inert-gas-condensation technique.

**Chemical Synthesis:** Chemical precipitation, chemical vapor deposition, micro-emulsions, spray pyrolysis, thermal spraying.

### B. Production of Nanofluids

There are mainly two methods of nanofluid production, namely, two-step technique and one-step technique. In the two-step technique, the first step is the production of nanoparticles and the second step is the dispersion of the nanoparticles in a base fluid. Two-step technique is advantageous when mass production of nanofluids is

A. PG student in Department of Mechanical Engineering (ME Thermal) Gandhinagar Institute of Technology Gandhinagar, India (email- milindpatel@rocketmail.com)

B. PG student in Department of Mechanical Engineering (ME Thermal) Gandhinagar Institute of Technology Gandhinagar, India (email- patel.ankit27789@gmail.com)

considered, because at present, nanoparticles can be produced in large quantities by utilizing the technique of inert gas condensation [6]. The main disadvantage of the two-step technique is that the nanoparticles form clusters during the preparation of the nanofluid which prevents the proper dispersion of nanoparticles inside the base fluid [5].

One-step technique combines the production of nanoparticles and dispersion of nanoparticles in the base fluid into a single step. There are some variations of this technique. In one of the common methods, named direct evaporation one-step method, the nanofluid is produced by the solidification of the nanoparticles, which are initially gas phase, inside the base fluid [4]. The dispersion characteristics of nanofluids produced with one-step techniques are better than those produced with two-step technique [5]. The main drawback of one-step techniques is that they are not proper for mass production, which limits their commercialization [5].

#### *Two- Step Method*

Two-step method is the most widely used method for preparing nanofluids. Nanoparticles, nanofibers, nanotubes, or other nanomaterials used in this method are first produced as dry powders by chemical or physical methods. Then, the nanosized powder will be dispersed into a fluid in the second processing step with the help of intensive magnetic force agitation, ultrasonic agitation, high-shear mixing, homogenizing, and ball milling. Two-step method is the most economic method to produce nanofluids in large scale, because nanopowder synthesis techniques have already been scaled up to industrial production levels. Due to the high surface area and surface activity, nanoparticles have the tendency to aggregate. The important technique to enhance the stability of nanoparticles in fluids is the use of surfactants. However, the functionality of the surfactants under high temperature is also a big concern, especially for high-temperature applications.

Due to the difficulty in preparing stable nanofluids by two-step method, several advanced techniques are developed to produce nanofluids, including one-step method. In the following part, we will introduce one-step method in detail.

#### *One- Step Method*

To reduce the agglomeration of nanoparticles, Eastman et al. developed a one-step physical vapor condensation method to prepare Cu/ethylene glycol nanofluids [8]. The one-step process consists of simultaneously making and dispersing the particles in the fluid. In this method, the processes of drying, storage, transportation, and dispersion of nanoparticles are avoided, so the agglomeration of nanoparticles is minimized, and the stability of fluids is increased [7]. The one-step processes can prepare uniformly dispersed nanoparticles, and the particles can be stably suspended in the base fluid. The vacuum-SANSS (submerged arc nanoparticle synthesis system) is another efficient method to prepare nanofluids using different dielectric liquids [9], [10]. The different morphologies are mainly influenced and determined by

various thermal conductivity properties of the dielectric liquids. The nanoparticles prepared exhibit needle-like, polygonal, square, and circular morphological shapes. The method avoids the undesired particle aggregation fairly well.

One-step physical method cannot synthesize nanofluids in large scale, and the cost is also high, so the one-step chemical method is developing rapidly. Zhu et al. presented a novel one-step chemical method for preparing copper nanofluids by reducing  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  with  $\text{NaH}_2\text{PO}_2 \cdot \text{H}_2\text{O}$  in ethylene glycol under microwave irradiation [11]. Well-dispersed and stably suspended copper nanofluids were obtained. Mineral oil-based nanofluids containing silver nanoparticles with a narrow-size distribution were also prepared by this method [12]. The particles could be stabilized by Korantin, which coordinated to the silver particle surfaces via two oxygen atoms forming a dense layer around the particles. The silver nanoparticle suspensions were stable for about 1 month. Stable ethanol-based nanofluids containing silver nanoparticles could be prepared by microwave-assisted one-step method [13]. In the method, polyvinylpyrrolidone (PVP) was employed as the stabilizer of colloidal silver and reducing agent for silver in solution. The cationic surfactant octadecylamine (ODA) is also an efficient phase-transfer agent to synthesize silver colloids [14]. The phase transfer of the silver nanoparticles arises due to coupling of the silver nanoparticles with the ODA molecules present inorganic phase via either coordination bond formation or weak covalent interaction. Phase transfer method has been developed for preparing homogeneous and stable graphene oxide colloids. Graphene oxide nanosheets (GONs) were successfully transferred from water to n-octane after modification by Oley amine, and the schematic illustration of the phase transfer process is shown in Fig. 1 [15].

However, there are some disadvantages for one-step method. The most important one is that the residual reactants are left in the nanofluids due to incomplete reaction or stabilization. It is difficult to elucidate the nanoparticle effect without eliminating this impurity effect.

#### *Other Novel Methods*

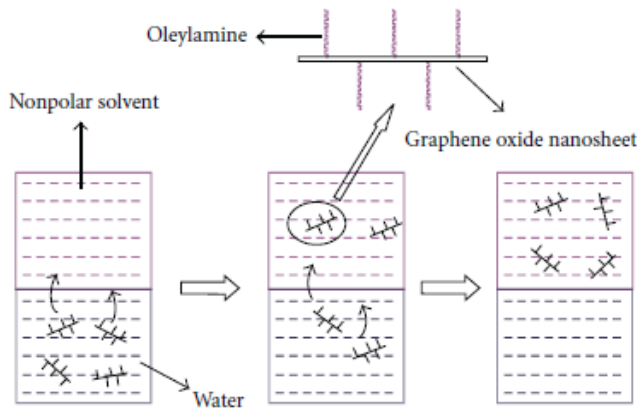


Fig. 1. Schematic illustration of the phase transfer process.

Wei et al. developed a continuous flow micro fluid in micro reactor to synthesize copper nanofluids. By this method, copper nanofluids can be continuously synthesized, and their microstructure and properties can be varied by adjusting parameters such as reactant concentration, flow rate, and additive. CuO nanofluids with high solid volume fraction (up to 10 vol. %) can be synthesized through a novel precursor transformation method with the help of ultrasonic and microwave irradiation [16]. The precursor  $\text{Cu}(\text{OH})_2$  is completely transformed to CuO nanoparticle in water under microwave irradiation. The ammonium citrate prevents the growth and aggregation of nanoparticles, resulting in a stable CuO aqueous nanofluid with higher thermal conductivity than those prepared by other dispersing methods. Phase-transfer method is also facile way to obtain monodisperse noble metal colloids [17]. In a water cyclohexane two-phase system, aqueous formaldehyde is transferred to cyclohexane phase via reaction with dodecylamine to form reductive intermediates in cyclohexane. The intermediates are capable of reducing silver or gold ions in aqueous solution to form dodecylamine-protected silver and gold nanoparticles in cyclohexane solution at room temperature. Feng et al. used the aqueous organic phase transfer method for preparing gold, silver, and platinum nanoparticles on the basis of the decrease of the PVP's solubility in water with the temperature increase [18]. Phase transfer method is also applied for preparing stable kerosene based  $\text{Fe}_3\text{O}_4$  nanofluids. Oleic acid is successfully grafted onto the surface of  $\text{Fe}_3\text{O}_4$  nanoparticles by chemisorbed mode, which lets  $\text{Fe}_3\text{O}_4$  nanoparticles have good compatibility with kerosene [19]. The  $\text{Fe}_3\text{O}_4$  nanofluids prepared by phase-transfer method do not show the previously reported "time dependence of the thermal conductivity characteristic". The preparation of nanofluids with control able microstructure is one of the key issues. It is well known that the properties of nanofluids strongly depend on the structure and shape of nanomaterials. The recent research shows that nanofluids synthesized by chemical solution method have both higher conductivity enhancement and better stability than those produced by the other methods

[20]. This method is distinguished from the others by its controllability. The nanofluid microstructure can be varied and manipulated by adjusting synthesis parameters such as temperature, acidity, ultrasonic and microwave irradiation, types and concentrations of reactants and additives, and the order in which the additives are added to the solution.

#### IV. THERMAL CONDUCTIVITY OF NANOFLUID

Studies regarding the thermal conductivity of nanofluids showed that high enhancements of thermal conductivity can be achieved by using nanofluids. It is possible to obtain thermal conductivity enhancements larger than 20% at a particle volume fraction smaller than 5% [21]-[23]. Such enhancement values exceed the predictions of theoretical models developed for suspensions with larger particles. This is considered as an indication of the presence of additional thermal transport enhancement mechanisms of nanofluids.

##### *Thermal Conductivity Measurement Methods*

In thermal conductivity measurements of nanofluids, the transient hot-wire technique is the most commonly used method [24]-[27]. A modified transient hot-wire method is required in the measurements, since nanofluids conduct electricity. The modification is made by insulating the wire. Some other methods such as steady-state parallel-plate technique, temperature oscillation technique, micro hot strip method, and optical beam deflection technique have also been utilized by some researchers [28]-[31].

#### V. EFFECTS OF SOME PARAMETERS ON THERMAL CONDUCTIVITY OF NANOFLUIDS

Experimental studies show that thermal conductivity of nanofluids depends on many factors such as particle volume fraction, particle material, particle size, particle shape, base fluid material, and temperature.

##### *A. Particle Volume Fraction*

Masuda et al. [3] measured the thermal conductivity of nanofluids containing  $\text{Al}_2\text{O}_3$  (13 nm),  $\text{SiO}_2$  (12 nm), and  $\text{TiO}_2$  (27 nm) nanoparticles (values in parentheses indicate the average particle diameter). This is the first experimental study regarding the thermal conductivity of nanofluids. Water was used as the base fluid and a two-step method was utilized for the preparation of nanofluids. An enhancement as high as 32.4% was observed for the effective thermal conductivity of 4.3 vol. %  $\text{Al}_2\text{O}_3$ /water nanofluid at  $31.85^\circ\text{C}$  (all percentage enhancement values are indicated according to the expression  $100(\text{k}_{\text{nf}} - \text{k}_f)/\text{k}_f$  throughout the discussion). It was found that thermal conductivity enhancement increases linearly with particle volume fraction. Lee et al. [32] studied the room temperature thermal conductivity of nanofluids by dispersing  $\text{Al}_2\text{O}_3$  (38.5 nm) and CuO (23.6 nm) nanoparticles, which were produced by gas condensation method, in water and ethylene glycol. Similar to the study of Masuda et al. [3], a

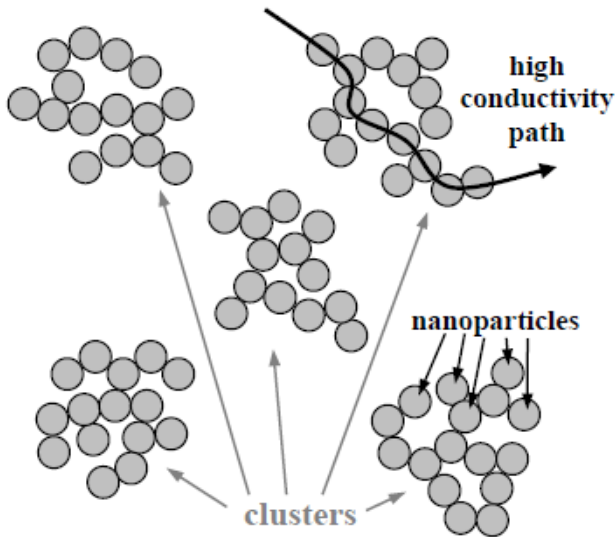


Fig. 2. Schematic illustration representing the clustering phenomenon [42]. High conductivity path results in fast transport of heat along large distances.

linear relationship was observed between thermal conductivity and particle volume fraction (thermal conductivity increases with particle volume fraction). Highest enhancement was 20%, which was observed for 4 vol. % CuO/ethylene glycol nanofluid. A similar study was performed by Wang et al. [28], who examined the thermal conductivity performance of nanofluids with Al<sub>2</sub>O<sub>3</sub> (28 nm) and CuO (23 nm) nanoparticles. For the case of 8 vol. % Al<sub>2</sub>O<sub>3</sub>/water nanofluid, thermal conductivity enhancement as high as 40% was achieved. For water- and ethylene glycol-based nanofluids, thermal conductivity ratio showed a linear relationship with particle volume fraction and the lines representing this relation were found to be coincident.

### B. Particle Material

Most of the studies show that particle material is an important parameter that affects the thermal conductivity of nanofluids. At first glance, it might be thought that the difference in the thermal conductivities of particle materials is the main reason of this effect. However, studies show that particle type may affect the thermal conductivity of nanofluids in other ways. For example, Lee et al. [32] considered the thermal conductivity of nanofluids with Al<sub>2</sub>O<sub>3</sub> and CuO nanoparticles as mentioned in the previous section and they found that nanofluids with CuO nanoparticles showed better enhancement when compared to the nanofluids prepared using Al<sub>2</sub>O<sub>3</sub> nanoparticles. It should be noted that Al<sub>2</sub>O<sub>3</sub>, as a material, has higher thermal conductivity than CuO. Therefore, thermal conductivity of particle material may not be the dominant parameter that determines the thermal conductivity of the nanofluid. According to the authors, the key factor is the fact that Al<sub>2</sub>O<sub>3</sub> nanoparticles formed relatively larger clusters when compared to CuO nanoparticles. That might be an explanation if the main

mechanism of thermal conductivity enhancement is accepted to be the Brownian motion of nanoparticles, since the effect of Brownian motion diminishes with increasing particle size. However, it should also be noted that there are some studies that consider the clustering of nanoparticles as a thermal conductivity enhancement mechanism. Another study that considers the effect of nanoparticle type was made by Chopkar et al. [34]. They dispersed Al<sub>2</sub>Cu and Ag<sub>2</sub>Al nanoparticles into water and ethylene glycol. 1 vol. % oleic acid was added as the surfactant. Measurements were made at room temperature. It was found that Ag<sub>2</sub>Al nanoparticles enhanced thermal conductivity slightly more when compared to Al<sub>2</sub>Cu nanoparticles. According to the authors, this is due to the fact that the thermal conductivity of Ag<sub>2</sub>Al is higher when compared to Al<sub>2</sub>Cu.

### C. Base Fluid

According to the conventional thermal conductivity models such as the Maxwell model [35], as the base fluid thermal conductivity of a mixture decreases, the thermal conductivity ratio (thermal conductivity of nanofluid (knf) divided by the thermal conductivity of base fluid (kf)) increases. When it comes to nanofluids, the situation is more complicated due to the fact that the viscosity of the base fluid affects the Brownian motion of nanoparticles and that in turn affects the thermal conductivity of the nanofluid [36]. Moreover, Lee [37] examined the effect of electric double layer forming around nanoparticles on the thermal conductivity of nanofluids and showed that the thermal conductivity and thickness of the layer depends on the base fluid. It is difficult to determine the quantitative effects of these factors completely. Therefore, systematic experiments are required that will show the effect of base fluid on the thermal conductivity of nanofluids. Some experimental studies made in this area are summarized below.

In the previously mentioned study of Wang et al. [28], Al<sub>2</sub>O<sub>3</sub> and CuO nanoparticles were used to prepare nanofluids with several base fluids; water, ethylene glycol, vacuum pump fluid, and engine oil. With Al<sub>2</sub>O<sub>3</sub> nanoparticles, the highest thermal conductivity ratio was observed when ethylene glycol was used as the base fluid. Engine oil showed somewhat lower thermal conductivity ratios than ethylene glycol. Water and pump fluid showed even smaller ratios, respectively. With CuO nanoparticles, only ethylene glycol- and water-based nanofluids were prepared and it is interesting to note that they showed exactly the same thermal conductivity ratios for the same particle volume fraction. The effect of the base fluid on the thermal conductivity of nanofluids was also analyzed by Xie et al. [38]. Nanofluids with Al<sub>2</sub>O<sub>3</sub> nanoparticles were prepared by using different base fluids; deionized water, glycerol, ethylene glycol, and pump oil. In addition, ethylene glycol-water and glycerol-water mixtures with different volume fractions were also used as base fluids and the variation of the thermal conductivity ratio with thermal conductivity of the base fluid mixture was examined. It was seen that, thermal conductivity

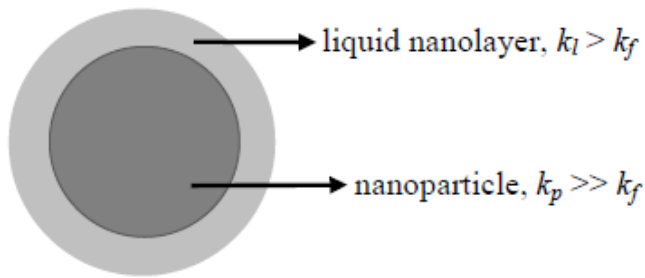


Fig. 3. Schematic illustration representing the liquid layering around nanoparticles.  $k_l$ ,  $k_f$ , and  $k_p$  are the thermal conductivity of nanolayer, base fluid, and nanoparticle, respectively.

ratio decreased with increasing thermal conductivity of the base fluid. Results were compared with a theoretical analysis made by Hasselman and Johnson [39]. Theoretical results were found to be nearly independent of the thermal conductivity of the base fluid, being contrary to the experimental data. However, it should be noted that these experimental results are in agreement with the Maxwell model [35] qualitatively. Chopkar et al. [34] also analyzed the effect of base fluid by comparing water and ethylene glycol. Al<sub>2</sub>Cu and Ag<sub>2</sub>Al nanoparticles were used in the study and it was found that water-based nanofluids showed a higher thermal conductivity ratio. It should be noted that more than 100% enhancement was obtained for the 2.0 vol. % Ag<sub>2</sub>Al (30 nm)/water nanofluid.

#### D. Particle Size

Particle size is another important parameter of thermal conductivity of nanofluids. It is possible to produce nanoparticles of various sizes, generally ranging between 5 and 100 nm. Chopkar et al. [41] prepared nanofluids by dispersing Al<sub>70</sub>Cu<sub>30</sub> nanoparticles into ethylene glycol. Nanoparticles were obtained by mechanical alloying. By transmission electron microscopy, they illustrated the fact that there is no significant clustering in the samples. They varied the particle size between 9 and 83 nm and they showed that thermal conductivity enhancement decreases with increasing particle size. For 0.5 vol. % nanofluid, thermal conductivity enhancement decreased from 38 to 3% by increasing the particle size from 9 to 83 nm. The general trend in the experimental data is that the thermal conductivity of nanofluids increases with decreasing particle size.

#### E. Particle Shape

There are mainly two particle shapes used in nanofluid research; spherical particles and cylindrical particles. Cylindrical particles usually have a large length-to-diameter ratio. The thermal conductivity of SiC/distilled water and SiC/ethylene glycol nanofluids were investigated by Xie et al. [43]. Two types of nanoparticles were used for the preparation of nanofluids; spherical particles with 26 nm average diameter and cylindrical particles with 600 nm average diameter. It was found that 4.2 vol. % water-based nanofluid with spherical particles had a thermal conductivity

enhancement of 15.8%, whereas 4 vol. % nanofluid with cylindrical particles had a thermal conductivity enhancement of 22.9%. The authors compared the results with the Hamilton and Crosser model [40]. It was noted that Hamilton and Crosser model was successful in predicting the enhancement in cylindrical particles, whereas it underestimated the values associated with nanofluids with spherical particles. Another study related to the particle shape was made by Murshed et al. [33]. They measured the thermal conductivity of TiO<sub>2</sub>/deionized water nanofluid.

In addition to these experimental results, the fact that nanofluids with carbon nanotubes (which are cylindrical in shape) generally show greater thermal conductivity enhancement than nanofluids with spherical particles should also be considered. As a result, one can conclude that cylindrical nanoparticles provide higher thermal conductivity enhancement than spherical particles. One of the possible reasons of this is the rapid heat transport along relatively larger distances in cylindrical particles since cylindrical particles usually have lengths on the order of micrometers. However, it should be noted that nanofluids with cylindrical particles usually have much larger viscosities than those with spherical nanoparticles [44]. As a result, the associated increase in pumping power is large and this reduces the feasibility of usage of nanofluids with cylindrical particles.

#### F. Temperature

In conventional suspensions of solid particles (with sizes on the order of millimeters or micrometers) in liquids, thermal conductivity of the mixture depends on temperature only due to the dependence of thermal conductivity of base liquid and solid particles on temperature. However, in case of nanofluids, change of temperature affects the Brownian motion of nanoparticles and clustering of nanoparticles [45], which results in dramatic changes of thermal conductivity of nanofluids with temperature. Masuda et al. [3] measured the thermal conductivity of water-based nanofluids containing Al<sub>2</sub>O<sub>3</sub>, SiO<sub>2</sub>, and TiO<sub>2</sub> nanoparticles at different temperatures. Thermal conductivity ratio decreased with increasing temperature, which is contradictory to many findings in the literature. The temperature dependence of the thermal conductivity of Al<sub>2</sub>O<sub>3</sub> (38.4 nm)/water and CuO (28.6 nm)/water nanofluids was studied by Das et al. [22]. Thermal diffusivity was measured by using a temperature oscillation technique and then thermal conductivity was calculated. Several measurements were made at different temperatures varying between 21 and 51°C. It was seen that for 1 vol. % Al<sub>2</sub>O<sub>3</sub>/water nanofluid, thermal conductivity enhancement increased from 2% at 21°C to 10.8% at 51°C. Temperature dependence of 4 vol.% Al<sub>2</sub>O<sub>3</sub> nanofluid was much more significant. From 21 to 51°C, enhancement increased from 9.4 to 24.3%. A linear relationship between thermal conductivity ratio and temperature was observed at both 1 and 4 vol. % cases.

Li and Peterson [23] also investigated the effect of temperature on thermal conductivity of CuO (29 nm)/water

and Al<sub>2</sub>O<sub>3</sub> (36 nm)/water nanofluids. For both nanofluid types, it was observed that at a constant particle volume fraction thermal conductivity ratio increased with temperature. In addition, it was noted that for Al<sub>2</sub>O<sub>3</sub>/water nanofluid, the dependence of thermal conductivity ratio on particle volume fraction became more pronounced with increasing temperature.

### G. Clustering

Clustering is the formation of larger particles through aggregation of nanoparticles. Clustering effect is always present in nanofluids and it is an effective parameter in thermal conductivity. Evans et al. [48] proposed that clustering can result in fast transport of heat along relatively large distances since heat can be conducted much faster by solid particles when compared to liquid matrix. This phenomenon is illustrated schematically in Fig. 2.

### H. pH Value

The number of studies regarding the pH value of nanofluids is limited when compared to the studies regarding the other parameters. Xie et al. [46] measured the thermal conductivity of nanofluids, which are prepared by dispersing Al<sub>2</sub>O<sub>3</sub> nanoparticles into water, ethylene glycol, and pump oil. They reported significant decrease in thermal conductivity ratio with increasing pH values. It was also observed that the rate of change of thermal conductivity with particle volume fraction was dependent on pH value. Thermal conductivity enhancement of 5 vol. % Al<sub>2</sub>O<sub>3</sub>/water nanofluid was 23% when pH is equal to 2.0 and it became 19% when pH is equal to 11.5. Wang et al. [47] also investigated the effect of pH on the thermal conductivity of nanofluids. They considered Cu/water and Al<sub>2</sub>O<sub>3</sub>/water nanofluids, as the dispersant, sodium dodecylbenzene sulfonate, were added to the samples. They obtained optimum values of pH (approximately 8.0 for Al<sub>2</sub>O<sub>3</sub>/water and 9.5 for Cu/water nanofluids) for maximum thermal conductivity enhancement. It should also be noted that the thermal conductivity of base fluid does not change significantly with pH. The authors related the observed phenomenon to the fact that at the optimum value of pH, surface charge of nanoparticles increases, which creates repulsive forces between nanoparticles. As a result of this effect, severe clustering of nanoparticles is prevented (excessive clustering may result in sedimentation, which decreases thermal conductivity enhancement).

### I. Liquid Layering around Nanoparticles

A recent study showed that liquid molecules form layered structures around solid surfaces [49] and it is expected that those nanolayers have a larger effective thermal conductivity than the liquid matrix [50]. As a result of this observation, the layered structures that form around nanoparticles are proposed to be responsible for the thermal conductivity enhancement of nanofluids [50]. This phenomenon is illustrated schematically in Fig. 3.

## VI. CONCLUSION & FUTURE WORK

Many interesting properties of nanofluids have been reported in the past decades. We can conclude that different nanofluid preparation methods have different merits & demerits, also properties of nanofluids depend on many factors. All these factors are important to study the behavior of nanofluid.

The research about the effects of clustering, pH value, and ultrasonic vibration on thermal conductivity is very limited and further research is required regarding the effects of these parameters.

Similar to the case of convective heat transfer of nanofluids is also dependent on many parameters such as particle volume fraction, particle size, particle material, temperature, and base fluid type. Detailed experimental investigation of the effects of most of these parameters on heat transfer has not been performed yet. Systematic studies about these aspects of nanofluid heat transfer will provide valuable information for the optimization of heat transfer enhancement obtained with nanofluids.

### Nomenclature

knf	Thermal conductivity of nanofluid
kf	Thermal conductivity of base fluid
kl	Thermal conductivity of nanolayer
kp	Thermal conductivity of nanoparticle

## REFERENCES

- [1] B. Wang, L. Zhou, and X. Peng, "A Fractal Model for Predicting the Effective Thermal Conductivity of Liquid with Suspension of Nanoparticles," *International Journal of Heat and Mass Transfer*, 2003, 46(14), pp. 2665-2672.
- [2] P. Keblinski, S. R. Phillpot, S. U. S. Choi, and J. A. Eastman, "Mechanisms of Heat Flow in Suspensions of Nano-Sized Particles (Nanofluids)," *International Journal of Heat and Mass Transfer*, 2002, 45(4), pp. 855-863
- [3] H. Masuda, A. Ebata, K. Teramae, and N. Hishinuma, "Alteration of Thermal Conductivity and Viscosity of Liquid by Dispersing Ultra-Fine Particles (Dispersion of  $\gamma$ -Al<sub>2</sub>O<sub>3</sub>, SiO<sub>2</sub>, and TiO<sub>2</sub> Ultra-Fine Particles)," *Netsu Bussei*, 1993, 4(4), pp. 227-233.
- [4] J. A. Eastman, S. U. S. Choi, S. Li, W. Yu, and L. J. Thompson, "Anomalous Increased Effective Thermal Conductivities of Ethylene Glycol-Based Nanofluids Containing Copper Nanoparticles," *Applied Physics Letter*, 2001, 78(6), pp. 718-720.
- [5] W. Yu, D. M. France, J. L. Routbort, and S. U. S. Choi, "Review and Comparison of Nanofluid Thermal Conductivity and Heat Transfer Enhancements," *Heat Transfer Engineering*, 2008, 29(5), pp. 432-460.
- [6] J. M. Romano, J. C. Parker, and Q. B. Ford, "Application Opportunities for Nanoparticles Made from the Condensation of Physical Vapors," *Advances Powder Metallurgy & Particulate Materials*, 1997, pp. 12-13.
- [7] Y. Li, J. Zhou, S. Tung, E. Schneider, and S. Xi, "A review on development of nanofluid preparation and characterization," *Powder Technology*, 2009, vol. 196, no. 2, pp. 89-101.
- [8] J. A. Eastman, S. U. S. Choi, S. Li, W. Yu, and L. J. Thompson, "Anomalous increased effective thermal conductivities of ethylene glycol-based nanofluids containing copper nanoparticles," *Applied Physics Letters*, 2001, vol. 78, no. 6, pp. 718-720.
- [9] C. H. Lo, T. T. Tsung, and L. C. Chen, "Shape-controlled synthesis of Cu-based nanofluid using submerged arc nanoparticle synthesis system (SANSS)," *Journal of Crystal Growth*, 2005, vol.277, no. 1-4, pp. 636-642.



- [10] C. H. Lo, T. T. Tsung, L. C. Chen, C. H. Su, and H. M. Lin, "Fabrication of copper oxide nanofluid using submerged arc nanoparticle synthesis system (SANS),", *Journal of Nanoparticle Research*, 2005, vol. 7, no. 2-3, pp. 313–320.
- [11] H. T. Zhu, Y. S. Lin, and Y. S. Yin, "A novel one-step chemical method for preparation of copper nanofluids," *Journal of Colloid and Interface Science*, 2004, vol. 277, no. 1, pp. 100–103.
- [12] H. B'onnemann, S. S. Botha, B. Bladergroen, and V. M. Linkov, "Monodisperse copper and silver nanocolloids suitable for heat conductive fluids," *Applied Organ metallic Chemistry*, 2005, vol. 19, no. 6, pp. 768–773.
- [13] A. K. Singh and V. S. Raykar, "Microwave synthesis of silver nanofluids with polyvinylpyrrolidone (PVP) and their transport properties," *Colloid and Polymer Science*, 2008, vol. 286, no. 14-15, pp. 1667–1673.
- [14] A. Kumar, H. Joshi, R. Pasricha, A. B. Mandale, and M. Sastry, "Phase transfer of silver nanoparticles from aqueous to organic solutions using fatty amine molecules," *Journal of Colloid and Interface Science*, 2003, vol. 264, no. 2, pp. 396–401.
- [15] W. Yu, H. Xie, X. Wang, and X. Wang, "Highly efficient method for preparing homogeneous and stable colloids containing graphene oxide," *Nanoscale Research Letters*, 2011, vol. 6, p. 47.
- [16] H. T. Zhu, C. Y. Zhang, Y. M. Tang, and J. X. Wang, "Novel synthesis and thermal conductivity of CuO nanofluid," *Journal of Physical Chemistry C*, 2007, vol. 111, no. 4, pp. 1646–1650.
- [17] Y. Chen and X. Wang, "Novel phase-transfer preparation of monodisperse silver and gold nanoparticles at room temperature," *Materials Letters*, 2008, vol. 62, no. 15, pp. 2215–2218.
- [18] X. Feng, H. Ma, S. Huang et al., "Aqueous-organic phase transfer of highly stable gold, silver, and platinum nanoparticles and new route for fabrication of gold nanofilms at the oil/water interface and on solid supports," *Journal of Physical Chemistry B*, 2006, vol. 110, no. 25, pp. 12311–12317.
- [19] W. Yu, H. Xie, L. Chen, and Y. Li, "Enhancement of thermal conductivity of kerosene-based Fe<sub>3</sub>O<sub>4</sub> nanofluids prepared via phase-transfer method," *Colloids and Surfaces A*, 2010, vol. 355, no. 1–3, pp. 109–113.
- [20] L. Wang and J. Fan, "Nanofluids research: key issues", *Nanoscale Research Letters*, 2010, vol. 5, no. 8, pp. 1241–1252.
- [21] C. H. Chon, K. D. Kihm, S. P. Lee, and S. U. S. Choi, "Empirical Correlation Finding the Role of Temperature and Particle Size for Nanofluid (Al<sub>2</sub>O<sub>3</sub>) Thermal Conductivity Enhancement," *Applied Physics Letter*, 2005, 87(15), 153107.
- [22] S. K. Das, N. Putra, P. Thiesen, and W. Roetzel, "Temperature Dependence of Thermal Conductivity Enhancement for Nanofluids," *Journal of Heat Transfer*, 2003, 125(4), pp. 567-574.
- [23] C. H. Li, and G. P. Peterson, "Experimental Investigation of Temperature and Volume Fraction Variations on the Effective Thermal Conductivity of Nanoparticle Suspensions (Nanofluids)," *Journal of Applied Physics*, 2006, 99(8), 084314.
- [24] Y. Ding, H. Alias, D. Wen, and R. A. Williams, "Heat Transfer of Aqueous Suspensions of Carbon Nanotubes (CNT Nanofluids)," *International Journal of Heat and Mass Transfer*, 2006, 49(1-2), pp. 240-250.
- [25] H. T. Zhu, C. Y. Zhang, Y. M. Tang, and J. X. Wang, "Novel Synthesis and Thermal Conductivity of CuO Nanofluid," *Journal of Physical Chemistry C*, 2007, 111(4), pp. 1646-1650.
- [26] T. Hong, H. Yang, and C. J. Choi, "Study of the Enhanced Thermal Conductivity of Fe Nanofluids," *Journal of Applied Physics*, 2005, 97(6), 064311-4.
- [27] M. Beck, Y. Yuan, P. Warrier, and A. Teja, "The Effect of Particle Size on the Thermal Conductivity of Alumina Nanofluids," *Journal of Nanoparticle Research*, 2009, 11(5), pp. 1129-1136.
- [28] X. Wang, S. U. S. Choi, and X. Xu, "Thermal Conductivity of Nanoparticle - Fluid Mixture," *Journal of Thermophysics and Heat Transfer*, 1999, 13(4), pp. 474-480.
- [29] W. Czarnetzki, and W. Roetzel, "Temperature Oscillation Techniques for Simultaneous Measurement of Thermal Diffusivity and Conductivity," *International Journal of Thermophysics*, 1995, 16(2), pp. 413-422.
- [30] Y. S. Ju, J. Kim, and M. Hung, "Experimental Study of Heat Conduction in Aqueous Suspensions of Aluminum Oxide Nanoparticles," *Journal of Heat Transfer*, 2008, 130(9), 092403-6.
- [31] S. A. Putnam, D. G. Cahill, P. V. Braun, Z. Ge, and R. G. Shimmin, "Thermal Conductivity of Nanoparticle Suspensions," *Journal of Applied Physics*, 2006, 99(8), 084308-6.
- [32] S. Lee, S. U. S. Choi, S. Li, and J. A. Eastman, "Measuring Thermal Conductivity of Fluids Containing Oxide Nanoparticles," *Journal of Heat Transfer*, 1999, 121(2), pp. 280-289.
- [33] S. Murshed, K. Leong, and C. Yang, "Enhanced Thermal Conductivity of TiO<sub>2</sub>/Water Based Nanofluids," *International Journal of Thermal Sciences*, 2005, 44(4), pp. 367-373.
- [34] M. Chopkar, S. Sudarshan, P. Das, and I. Manna, "Effect of Particle Size on Thermal Conductivity of Nanofluid," *Metallurgical and Materials Transactions A*, 2008, 39(7), pp. 1535-1542.
- [35] J. C. Maxwell, *A Treatise on Electricity and Magnetism*, Clarendon Press, Oxford, 1873.
- [36] Y. Xuan, Q. Li, and W. Hu, "Aggregation Structure and Thermal Conductivity of Nanofluids," *AIChE Journal*, 2003, 49(4), pp. 1038-1043.
- [37] D. Lee, "Thermophysical Properties of Interfacial Layer in Nanofluids," *Langmuir*, 2007, 23(11), pp. 6011-6018.
- [38] H. Xie, J. Wang, T. Xi, Y. Liu, and F. Ai, "Dependence of the Thermal Conductivity of Nanoparticle-Fluid Mixture on the Base Fluid," *Journal of Materials Science Letters*, 2002, 21(19), pp. 1469-1471.
- [39] D. Hasselman, and L. F. Johnson, "Effective Thermal Conductivity of Composites with Interfacial Thermal Barrier Resistance," *Journal of Composite Materials*, 1987, 21(6), pp. 508-515.
- [40] R. L. Hamilton, and O. K. Crosser, "Thermal Conductivity of Heterogeneous Two-Component Systems," *Industrial & Engineering Chemistry Fundamentals*, 1962, 1(3), pp. 187-191.
- [41] M. Chopkar, P. K. Das, and I. Manna, "Synthesis and Characterization of Nanofluid for Advanced Heat Transfer Applications," *Scripta Materialia*, 2006, 55(6), pp. 549-552.
- [42] R. Prasher, P. E. Phelan, and P. Bhattacharya, "Effect of Aggregation Kinetics on the Thermal Conductivity of Nanoscale Colloidal Solutions (Nanofluid)," *Nano Letters*, 2006, 6(7), pp. 1529-1534.
- [43] H. Xie, J. Wang, T. Xi, and Y. Liu, "Thermal Conductivity of Suspensions Containing Nanosized SiC Particles," *International Journal of Thermophysics*, 2002, 23(2), pp. 571-580.
- [44] E. V. Timofeeva, J. L. Routbort, and D. Singh, "Particle Shape Effects on Thermophysical Properties of Alumina Nanofluids," *Journal of Applied Physics*, 2009, 106(1), 014304-10.
- [45] C. H. Li, W. Williams, J. Buongiorno, L. Hu, and G. P. Peterson, "Transient and Steady-State Experimental Comparison Study of Effective Thermal Conductivity of Al<sub>2</sub>O<sub>3</sub>/Water Nanofluids," *Journal of Heat Transfer*, 2008, 130(4), 042407-7.
- [46] H. Xie, J. Wang, T. Xi, Y. Liu, F. Ai, and Q. Wu, "Thermal Conductivity Enhancement of Suspensions Containing Nanosized Alumina Particles," *Journal of Applied Physics*, 2002, 91(7), pp. 4568-4572.
- [47] X. Wang, D. Zhu, and S. Yang, "Investigation of pH and SDBS on Enhancement of Thermal Conductivity in Nanofluids," *Chemical Physics Letter*, 2009, 470(1-3), pp. 107-111.
- [48] W. Evans, R. Prasher, J. Fish, P. Meakin, P. Phelan, and P. Keblinski, "Effect of Aggregation and Interfacial Thermal Resistance on Thermal Conductivity of Nanocomposites and Colloidal Nanofluids," *International Journal of Heat and Mass Transfer*, 2008, 51(5-6), pp. 1431-1438.
- [49] C. Yu, A. G. Richter, A. Datta, M. K. Durbin, and P. Dutta, "Observation of Molecular Layering in Thin Liquid Films Using X-ray Reflectivity," *Physical Review Letters*, 1999, 82(11), pp. 2326-2329.
- [50] W. Yu, and S. U. S. Choi, "The Role of Interfacial Layers in the Enhanced Thermal Conductivity of Nanofluids: A Renovated Maxwell Model," *Journal of Nanoparticle Research*, 2003, 5(1), pp. 167-171.

# Steam Jet Refrigeration Systems powered by waste heat of thermal power plant With Water and R-11 as Refrigerant

A. Ashishnath D. Nathji

**Abstract**— In this crucial time when everyone is feeling to conserve the energy, and to minimize the heat loss to the environment, every step taken in the way to use waste heat energy is admirable. We cannot minimize our day to day requirement of energy (heat) but we can minimize the loss. We can also minimize the loss by using Waste heat recovery phenomena. Waste heat recovery phenomena are widely used now days. Waste heat is defined as energy which is rejected to environment after getting required work from cycle. Similarly in thermal power plant after work done by the turbine on the steam (heat) we generally reject that steam (heat) in to environment. Of course now a days super heater, economizer are already employed to use some part of steam (heat), but by Steam jet air refrigeration system we can use some part of that steam (heat) to get cooling effect also. Steam jet refrigeration systems can be designed as per the availability of steam pressure. Low pressure steam can also be used, which proves to be very economical. Since water is the cooling agent there is no pollution of the environment which Eliminates ozone pollution, so it is also proves it as eco friendly system.

**Index terms**- conserve, Waste heat recovery, ozone pollution, ecofriendly system

## I. INTRODUCTION

In steam jet refrigeration systems, water can be used as the refrigerant. Like air, it is perfectly safe. These systems were applied successfully to refrigeration in the early years of this century. At low temperatures the saturation pressures are low (0.008129 bar at 4°C) and the specific volumes are high (157.3 m<sup>3</sup>/kg at 4°C). The temperatures that can be attained using water as a refrigerant are not low enough for most refrigeration applications but are in the range which may satisfy air conditioning, cooling, or chilling requirements. Also, these systems are used in some chemical industries for several processes, e.g. the removal of paraffin wax from lubricating oils. Note that steam jet refrigeration systems are not used when temperatures below 5°C are required. The main advantages of this system are the utilization of mostly low-grade energy and relatively small amounts of shaft work.

A. PG student, Department of Mechanical Engineering (ME Thermal) (Student), Gandhinagar Institute of Technology, Gandhinagar, India ashishnathji@gmail.com

Steam jet refrigeration systems use steam ejectors to reduce the pressure in a tank containing the return water from a

Chilled water system. The steam jet ejector utilizes the energy of a fast-moving jet of steam to capture the flash tank vapor and compress it. Flashing a portion of the water in the tank reduces the liquid temperature. Figure given below presents a schematic arrangement of a steam jet refrigeration system for water cooling. In the system shown, high-pressure steam expands while flowing through the nozzle 1. The expansion causes a drop in pressure and an enormous increase in velocity. Due to the high velocity, flash vapor from the tank 2 is drawn into the swiftly moving steam and the mixture enters the diffuser 3. The velocity is gradually reduced in the diffuser but the pressure of the steam at the condenser 4 is increased 5-10 times more than that at the entrance of the diffuser (e.g. from 0.01 bar to 0.07 bar).

This pressure value corresponds to the condensing temperature of 40°C. This means that the mixture of high-pressure steam and the flash vapor may be liquefied in the condenser. The latent heat of condensation is transferred to the condenser water, which may be at 25 °C. The condensate 5 is pumped back to the boiler, from which it may again be

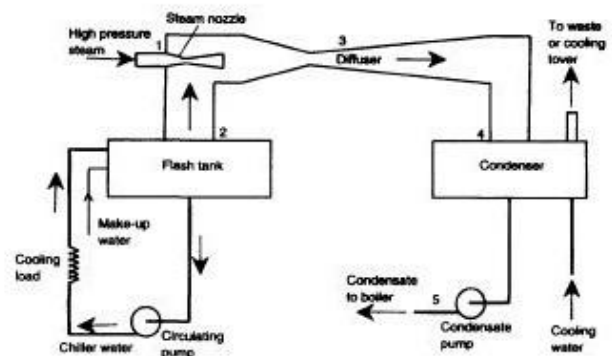


Fig. 1. Steam Jet Refrigeration system

vaporized at a high pressure. The evaporation of a relatively small amount of water in the flash tank (or flash cooler) reduces the temperature of the main body of water. The cooled water is then pumped as the refrigeration carrier to the cooling-load heat exchanger.

An ejector was invented by Sir Charles Parsons around 1901 for removing air from steam engine condensers. In about 1910, the ejector was used by Maurice Leblanc in the steam ejector refrigeration system. It experienced a wave of popularity during the early 1930s for air conditioning large buildings. Steam ejector refrigeration cycles were later

supplanted by systems using mechanical compressors. Since that time, development and refinement of ejector refrigeration systems have been almost at a standstill as most efforts have been concentrated on improving vapor compression cycles.

Furthermore, another typical gas-driven ejector is shown schematically in Figure. High-pressure primary fluid (P) enters the primary nozzle, through which it expands to produce a low-pressure region at the exit plane (1). The high-velocity primary stream draws and entrains the secondary fluid (S) into the mixing chamber. The combined streams are assumed to be completely mixed at the end of the mixing chamber (2) and the flow speed is supersonic. A normal shock wave is then produced within the mixing chamber's throat (3), creating a compression effect and the flow speed is reduced to a subsonic value. Further compression of the fluid is achieved as the mixed stream flows through the subsonic

refrigerant vapor is evolved in a boiler to produce the primary fluid for the ejector. The ejector draws vapor refrigerant from the evaporator as its secondary. This causes the refrigerant to evaporate at low pressure and produce useful refrigeration. The ejector exhausts the refrigerant vapor to the condenser where it is liquefied. The liquid refrigerant accumulated in the condenser is returned to the boiler via a pump whilst the remainder is expanded through a throttling valve to the evaporator, thus completing the cycle. As the working input required to circulate the fluid is typically less than 1 % of the heat supplied to the boiler, the COP may be defined as the ratio of evaporator refrigeration load to heat input to the boiler as follows:

$$COP = \frac{Q_L}{Q_B}$$

where  $Q_L$  is evaporator refrigeration load, kW; and  $Q_B$  is heat input to the boiler, kW.

Recently, scientist developed a new jet ejector refrigeration system using R-11 as the refrigerant. All vessels in the systems were constructed from galvanized steel. The boiler was designed to be electrically heated, two 4 kW electric heaters being located at the lower end. At its upper end, three baffle plates were welded to the vessel to prevent liquid droplets being carried over with the refrigerant vapor. The evaporator design was similar to that of the boiler. A single 3 kW electric heater was used to simulate a cooling load. A water-cooled plate type heat exchanger was used as a condenser. Cooling water was supplied at 32°C. The boiler was covered with a 40 mm thickness of glass wool with aluminum foil backing. The evaporator was covered with a 30 mm thickness of neoprene foam rubber. A diaphragm pump was used to circulate liquid refrigerant from the receiver tank to the boiler and the evaporator. The pump was driven by a variable speed 1/4 hp motor. One drawback of using the diaphragm pump is cavitations of the liquid refrigerant in the suction line due to pressure drop through an inlet check-valve. Therefore a small chiller was used to sub-cool the liquid R-11 before entering the pump. Figure given below shows a detailed schematic diagram of the experimental ejector. The nozzle was mounted on a threaded shaft, which allowed the position of the nozzle to be adjusted. Two different mixing chambers with throat diameter of 8 mm were used: in mixing chamber no.1, the mixing section is constant area duct; in mixing chamber no.2, the mixing section is convergent duct.

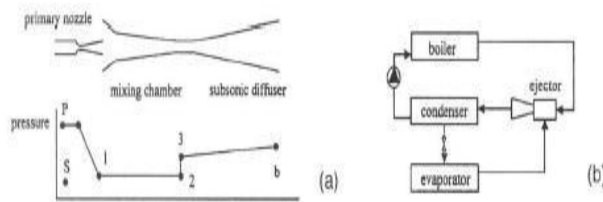


Fig. 2. Schematic of (a) a jet ejector and. (b) a simple jet ejector refrigeration system

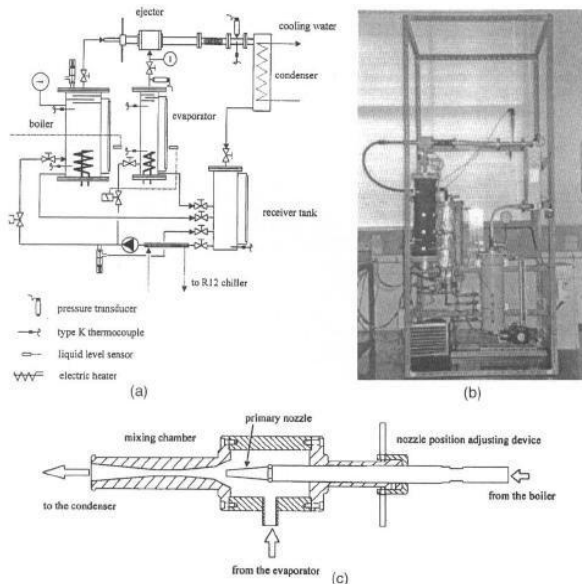


Fig. 3. (a) Schematic of the experimental ejector refrigerator. (b) Photograph of the experimental refrigerator. (c) The ejector used in the experimental set up

diffuser section (b).

Fig. 2 shows a schematic diagram of an ejector refrigeration cycle. It can be seen that a boiler, an ejector and a pump are used to replace the mechanical compressor of a conventional system. High-pressure and high-temperature

NOMENCLATURE

- $d$  Diameter ( $m$ )
- $Q_S$  Heat input, kW
- $Q_L$  Evaporator refrigeration Load, kW

SUBSCRIPTS

*f* Fluid  
*s* supplied  
*L* Load

REFERENCES

- [1] Arora & Domkandarwar, Refrigeration and Air conditioning.
- [2] H.K. Abdel-Aal, A.S. Al-Zakri, energy input for steam jet refrigeration systems, Department of Chemical Engineering, King Fahd University of Petroleum & Minerals, Dhahran 31261 Saudi Arabia
- [3] Ralph Mencke, Air conditioning by means of Steam Jet refrigeration.
- [4] Laboratories of Aphornratana et al. (2001)
- [5] [http://www.scribd.com/edrian\\_diaz/d/.../10-Steam-Jet-Refrigeration-System](http://www.scribd.com/edrian_diaz/d/.../10-Steam-Jet-Refrigeration-System).

# A Review on Behavior of Ductile Material under Thermal Cyclic Stress

A. Vinod P. Rajput, B. Dharmendrakumar R. Raval

**Abstract**— A review is presented of available information on the behavior of ductile material under condition of thermal stress and thermal cycle. For ductile materials, metallurgical effects on aging, corrosion, grain growth, residual stress due to thermal stress cycle when it is subjected to high temperature metallurgy are briefly discussed and illustrated from literature sources.

**Index Terms**—ductile material, thermal cycling model, metallurgical effects.

## I. INTRODUCTION

To The problem of thermal stress is of great importance in current high power energies. The present trend toward increasing temperatures has necessitated the use of refractory material capable of withstanding much higher temperature than normal engineering materials. One salient property of these materials is lack of ductility. For this reason, thermal stress is one of the most important design criteria in the application of these materials.

## II. THERMAL CYCLING MODEL

One of the mechanisms associated with the ultimate failure of ductile materials in thermal cycling is plastic flow. The role of the plastic flow process can be obtained by considering an idealized problem. The practical cases differ appreciably from the idealized problem.

In figure 1 is shown a bar fixed at its ends between two immovable plates so that the length of the bar must remain constant. This bar is assumed to be gradually cooled and heated between various temperature limits. Also shown in figure 1 is hypothetical stress-strain curve for material. It is assumed that material is assumed ideally plastic; that is, its stress-strain curve consists of straight line up to the yield stress and further yielding occurs at a constant stress. Thus, in figure 1 stresses is proportional to strain along the line OA

and further strain takes place at stress  $\sigma_A$  until rupture occurs when the strain is P. Let it be assumed that at the start of the process the bar is unstressed in the hot condition, and it is subsequently gradually cooled. The bar is taken as stress free in the heated condition in order to induce tensile stress during the first stage of the process. Because of constraint there must always be induced in the bar strain  $\alpha T_o$  equal to the thermal expansion, and the stress will depend on this strain  $\alpha T_o$  and on the stress-strain curve. As long as is less than the strain a stress is elastic and analysis would proceed just as if the material is brittle.

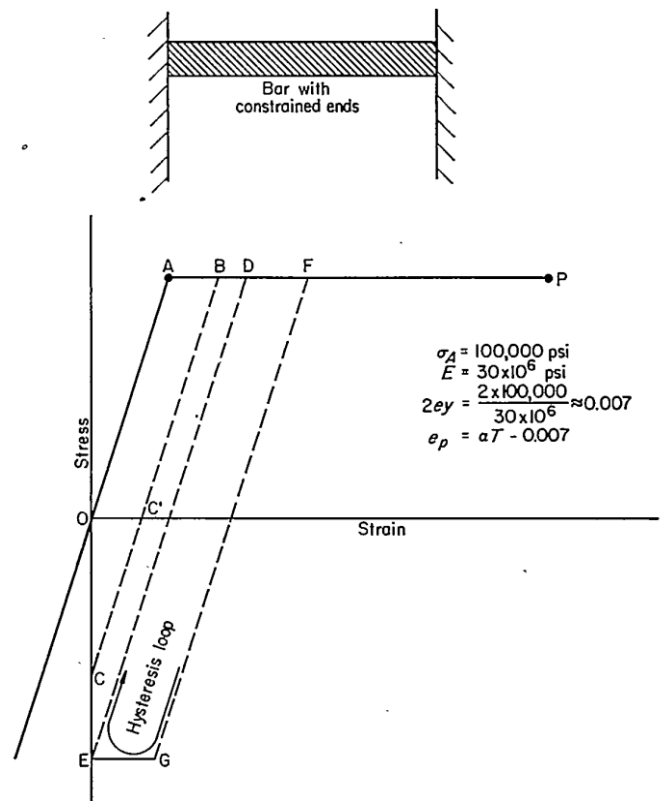


Fig. 1 Cyclic plastic strain in constrained bar due to heating and cooling [1]

Import If however, the cyclic temperature range is increased to induce the thermal strain equal to that at B, the stress developed will be the yield stress, and plastic flow of an amount AB will be introduced during the first cycle of temperature reduction. When the temperature is again raised

- A. PG student, Department of Mechanical Engineering (ME Thermal) (Student), Gandhinagar Institute of Technology, Gandhinagar, India (email- rajputvinodp@gmail.com)
- B. Department of Mechanical Engineering (ME Thermal) (Student), Gandhinagar Institute of Technology, Gandhinagar, India (email- dharmendraraval007@gmail.com)

to its initial value the stress will fall along a line BC'C. Sometime during the temperature increase the condition of the bar will be represented by C', where there is no stress but strain is not zero. When the initial temperature of the bar is finally reached the condition of the bar will be represented by point C- that is, zero strain and compressive stress OC, resulting from the fact that the free length of the bar had been increased by plastic flow by an amount AB.

If the temperature difference of cycling produces a thermal strain twice the elastic strain, that is, the strain at D is twice the strain at A, then cycling will occur between D and E, that is yield stress in tension and the yield stress in compression. (For simplicity, the yield stress in compression is here assumed to be equal to the yield stress in tension.). After the first cycle an indefinite number of cycle could be applied without further plastic flow.

Finally, consider an applied temperature difference such that the thermal strain is greater than the twice the elastic strain, for example point F. As the cooling cycle is applied the stress is first increased along the line OA and elastic strain occurs; then plastic flow occurs by an amount AF. As the specimen is subsequently heated it unloads elastically along line FG. At G, it still has not achieved its initial length because the strain is EG; and therefore as the temperature brought back to initial value, plastic flow at compression occurs by an amount EG. During the second cycle, the stress first changes from the compressive yield stress at E to the tensile yield stress at D and then further tensile plastic flow DF occurs. On the second unloading cycle the material proceed elastically from F to G and flow again in compression from G to E. Every cycle therefore induces in this bar plastic flow in tension of an amount DF and subsequently an equal amount of plastic flow. This alternate compressive and plastic flow ultimately leads to failure of material.[1]

The question is the number of cycles of this kind that this material will withstand. It has been found experimentally that the compressive plastic flow GE essentially improves the material so that even the sum of just the tensile components of plastic flow is greater than the ductility initially available in the conventional tensile test. That is, the summation of the DF values is greater than the strain DP. A simple assumption can be arrived at by a consideration of the data in a report by Sachs and co-workers [2]. In this investigation 24S-T aluminum alternately stretched and compressed by constant strain values and observation was made of the number of cycles that could be withstood before failure occurred.

Figure 2 shows typical results obtained in this investigation. The ductility remaining in the material after successive application of cycle of 12- percent tension and compression is shown on the vertical axis; the number of cycle application, on horizontal axis. The remaining ductility, after application of various numbers of cycles of alternating strain, was measured by subjecting a precycled specimen to a conventional tensile test in which ductility was measured. It is seen that the alloy had an initial ductility of 37 percent. After

a 12 strain in tension and a 12 – percent strain in compression were applied, the remaining ductility was reduced to 30 percent, a reduction less than that which would have been induced by just the 12-percent tension had not the compression followed. Similarly, it will be seen, each successive cycle reduced the remaining ductility by an amount less than 12 percent , and when the remaining ductility was less than 12 percent specimen failed upon the succeeding application of the strain cycle.

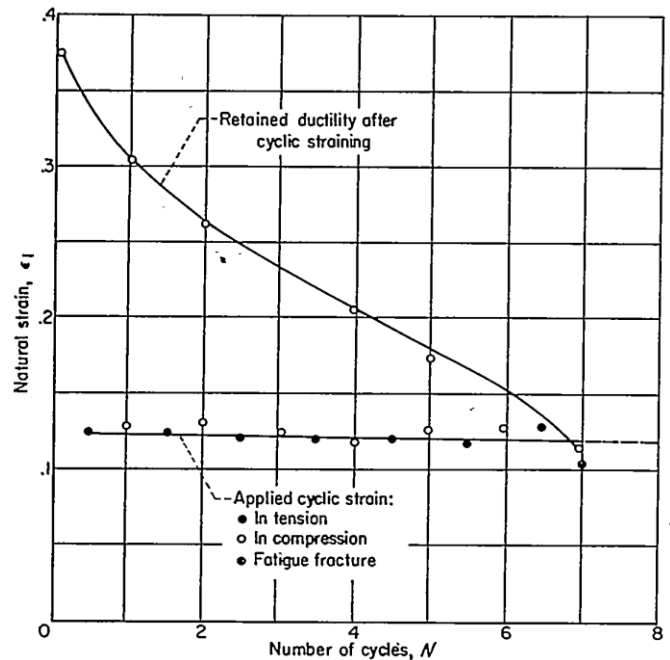


Fig. 2 Retained ductility of 248S-T alloy. Cyclic strain, ± 0.12 [2]

A plot of the data contained in ref [2] leads to the relation

$$N=K/\epsilon^n$$

(1)

- K proportionality constant
- N number of cycles of strain ± ε
- n exponent in the neighbourhood of 3
- ε plastic strain

The number of cycle is thus inversely proportional to approximately the cube of the strain per cycle.

A table can be prepared for the relative number of cycles that a material might be expected to withstand in the thermal fatigue type of test previously described. A material having the following properties will be considered (see figure 1): α= 10-6 per oF, E= 30×106 pounds per square inch and σ<sub>v</sub> =100,000 pounds per square inch. The elastic strain that this material can withstand is 0.0033 inch per inch. As previously described, no cyclic plastic flow will occur until the temperature is of the order of 700 °F . it is now assumed that the imposed cycling is between 0 and 1000 °F , 0 and 1200 °F, 0 and 1400 °F, 0 and 1600 °F, 0 and 1800 °F; and the number of cycles will be determined in each case. These

computations are as follows:

$$N_{1000} = K (1000 \times 10^{-5} - 0.007)^{-3}$$

$$N_{1200} = K (1200 \times 10^{-5} - 0.007)^{-3}$$

$$N_{1400} = K (1400 \times 10^{-5} - 0.007)^{-3}$$

$$N_{1600} = K (1600 \times 10^{-5} - 0.007)^{-3}$$

$$N_{1800} = K (1800 \times 10^{-5} - 0.007)^{-3}$$

Since these equations involve a constant K which is not known, only the relative number of each temperature can be determined. The results are shown in table 1. If the number of cycles for 0 to 1000 °F is taken as a base line, the number of cycles for the 0 to 1200 °F temperature range is reduced by factor of 4.7; temperature range from 0 to 1400 °F results in a reduction factor of 12.6; from 0 to 1600 °F, the reduction is by a factor of 27; and from 0 to 1800 °F, the reduction is by a factor of 49.

TABLE I  
RELATIVE REDUCTION OF THERMAL SHOCK CYCLES WITH INCREASE IN TEMPERATURE

Evaluation temperature °F	Relative reduction of thermal shock cycles at comparison temperature of-			
	1000 °F	1200 °F	1400 °F	1600 °F
1000	1.0			
1200	4.7	1.0		
1400	12.6	2.8	1.0	
1600	27	5.8	2.1	1.0
1800	49	9.7	3.9	1.6

Similarly, if the number of cycle at any other temperature ranges is used as references, the reduction in number of cycles for a higher temperature range is also shown in table.

### III. METALLURGICAL EFFECTS

During the thermal stress cycle, and between thermal stress cycles, the material is subjected to a complicated stress and temperature history that may cause metallurgical changes in in the microstructure. Mention will be made of only several of the metallurgical processes that can occur. The whole subject of high temperature metallurgy is really pertinent to this problem.

#### A. Aging

Probably the most important action that takes place during and in between thermal stress test is that of aging. Most high-temperature alloys in their use conditions are not in metallurgical equilibrium. In fact, it is because of their metastable condition that many alloy gain good high-temperature properties. If the material is maintained at high temperature, the tendency is toward rearrangement of the microstructure in the general direction of equilibrium. Thus, constituents that are in solid solution frequency tend to precipitate, and in so doing, they may drastically change the property of material. For example, they may precipitate n the gain boundaries and in this case they can reduce the ductility of the material, particularly in creep loading. This precipitation may occur

with or without the application of stress but generally stress tends to hasten the action. Finally, if the material becomes sufficiently embrittled, it may not be able to withstand even the small amount of plastic deformation required in a single thermal shock test. No careful investigation have, however, been made directly toward ascertaining the significance of precipitation in the thermal shock test. Its possible significance can be deduced by consideration of the studies that significance of precipitation in their types of mechanical tests.

Figure 3 shows some pertinent results originally published in the Russian literature and recently summarize by Sachs and Brown [7]. Plotted are the important strengths of three steels after they had been heated for certain periods of time at 932° F while in both the stressed and unstressed condition. The impact strength has been found to be sensitive to precipitation embitterment in cases when other types of tests miss detecting these effects completely. Note that heating alone of the Cr-Ni steel for about 100 hours reduced its impact strength tremendously. The presence of stress during heating cause this steel to embrittle even further, The Cr-Mo steel, on the hand, showed little embrittling effect due to the heating alone, with some embrittling resulting from the presence of stress. The stress was too law, however, for any conclusive statement about the significance of stress. The Cr-Ni-Mo showed an intermediate effect between the Cr-Ni and the Cr-Mo steel. Nickel is known to be associated with temper brittleness, a phenomenon believed but not proven, to be associated with grain boundary precipitation; hence, it is seen that precipitation may render and originally ductile material quite brittle. Of course, it would then become more susceptible to failure in thermal stress.

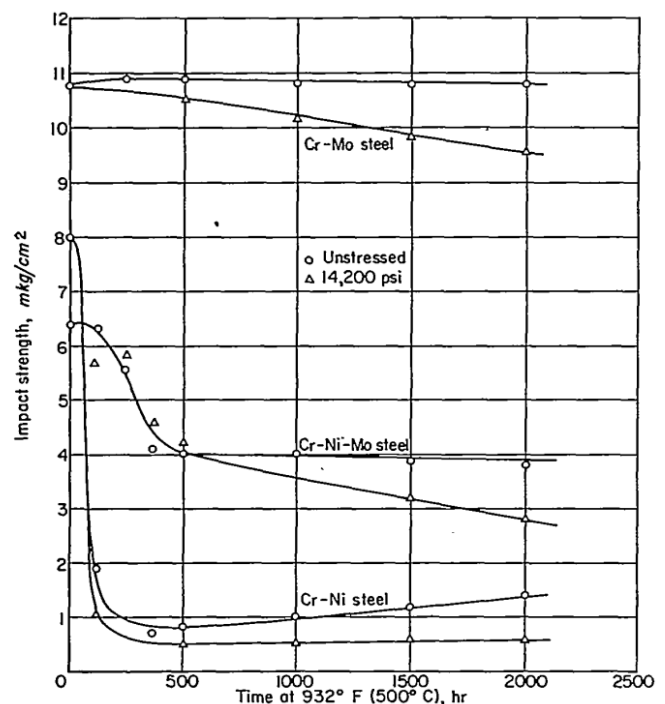


Fig. 3 Effect of heating with and without applied load on impact properties

[3]

The main lessons to be learned from figure 12 are the importance of aging even without stress in reducing impact strength and importance of the chemical composition on this effect. The aging need not take place during the thermal stress-it may occur during normal operation between thermal stress cycles. In laboratory thermal stress tests, it may occur during the heating process or during high-temperature soaking-the important factors time at temperature it allow the metallurgical reaction it take place. When the material is embrittled by all prior effects, thermal stress application will finally cause rupture. Since temperature is the most important variable affecting the aging process any excessive temperatures introduced in order to accelerate testing may produce foreign result.

As for the importance of chemical composition, the indication is that initial properties do not necessarily govern thermal stress behavior. Thus the Cr-Ni steel had better impact properties before aging than the Cr-Ni-Mo steel, but aging had more drastic effect on the former steel. It is not surprising, therefore, that some materials seem to perform better in thermal stress tests when their conventional properties do not indicate any reason for the superiority.

#### B. Corrosion.

Another process that may reduce thermal stress resistance is chemical attack, known otherwise as corrosion, oxidation, and so forth. The surface of the material is usually in contact with oxygen or other gas capable of chemical reaction with the material. At the high temperatures involved in stress testing, oxidation or other scales may form which are weak and brittle. Thermal shock testing then becomes a test, not of the original material but of resulting surface layer. Discontinuities formed at the surface layer either by cracking of the surface or by the disintegration of a corroded product act a source of stress concentration that induced and propagates further cracks within the body of the material. In some cases the corrosion consists not of the formation of a surface layer, but of diffusion into the body of the material. Hydrogen, for example, because of its small atomic dimension, readily diffuses into the grain boundaries of many materials, weakening them and rendering them less capable of withstanding thermal shock. The importance of intergranular attack is indicated by the fact that so many thermal shock failures are intergranular in nature.[7]

#### C. Grain growth

One of the effect of working is to render the material subjected to recrystallization. When grain are broken up, energy is stored in the slip planes and in the grain boundaries, and upon subsequent heating there is a tendency for the material to recrystallize in order to achieve a state of lower stored energy. In many cases the effect is to cause grain growth. Figure 4 shows some data obtained by Rush,

Freeman, and White [4] on the grain growth in the high-temperature alloy S-816 resulting from cold working. The plot shows the grain size resulting upon solution treatment at 2200° F as a function of the percent reduction per pass during the working process. It is seen that if the percent reduction per pass is approximately 0.75, which is a reasonable amount of plastic flow associated with a thermal shock cycle, the final grain size become A.S.T.M. 5. Grain size A.S.T.M. minus 2 is such a large grain size that an average turbine might contain only several such grains per cross section as compared with approximately one thousand grains per cross section for grain size A.S.T.M. 5. Materials with high grain size generally have low ductility and would not be expected to have high thermal shock resistance.

In general, a number of thermal stress tests described in the literature indicated that there was no effect on the grain size. However, observation has been limited, and attention is directed to this factor as only one of the many complicated metallurgical processes that can occur.

Of greater importance, perhaps, is that accelerated tests at excessively high temperatures may cause grain growth to occur while operating conditions may not. Hence, an accelerated test might give misleading results.

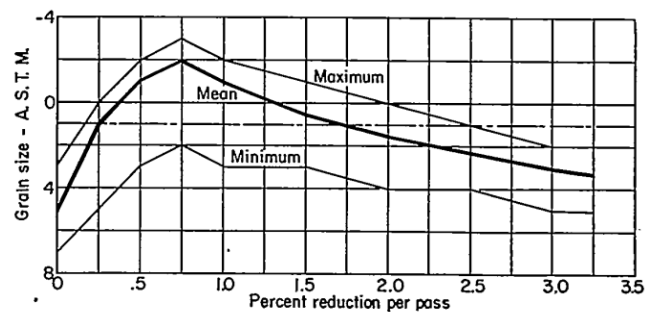


Fig. 4 Effect of prior critical deformation on grain size of s-816 alloy [4]

#### D. Residual stress

Another process that occurs in connection with thermal stress is the development and the relaxation of residual stress. As indicated in the previous simplified description of thermal cycling, the stress between the yield point in tension and the yield point in compression. However, during and between the thermal shock tests, the temperature constantly changes and residual stress introduction under one set of conditions may be caused to relax under temperature relief at another temperature. An example will later be given of relaxation of initial residual stress in connection with conventional fatigue tests at elevated temperatures.

The foregoing discussion represents only the more obvious of the complicated mechanisms that can occur in a high-temperature material during thermal stress testing. If the metallurgical factor is accepted as being probably the most important in the connection with thermal shock behavior of ductile materials, the significant conclusion to be reached relates to the importance of conducting tests to simulate actual operation. Metallurgical changes are very sensitive to



time and temperature. If, then an attempt is made to render the test more sever in order to expedite failure, misleading results can be obtained because of artificial metallurgical effects induced by the artificial test conditions.[4]

#### IV. CONCLUSION

It has been emphasized that for ductile materials the single-cycle criterion is not practical because it is rate that failure can be achieved in one cycle. The use of the multiple cycle criterion further complicates the already complex problem which involves metallurgical processes occurring during and in between thermal shock cycles. Some of these processes have been discussed and their potential rotation to thermal shock resistance indicated. In view of the importance of metallurgical processes, and of the importance of temperature influence on action of metallurgical processes, the danger of the conducting tests under artificial conditions become strikingly evident. Increasing temperature in order to accelerate failure may introduce spurious metallurgical effects, foreign to the behavior of the materials under their true working condition. Hence it is very important that the tests simulate the operating conditions irrespective of whether this entails a large number of tests before failure occur.

#### REFERENCES

- [1] S .S. Manson.: Behavior of Materials Under condition of Thermal Stress1.page no. 325-330, 1953.
- [2] S.I, Liu, Lynch, J.J., Ripling,E.J. and Sachs, G.: Low Cycle Fatigue of the Aluminum Alloy 24ST in Direct Stress. Tech Pub. No. 2388, Am. Inst. Mining and metallurgical Eng., Feb. 1948.
- [3] Sachs , Georges, and Brown, W.F.,Jr.: A survey of Embrittlement and Notch sensitivity of heat resisting steels. A.S.T.M. Spee. Tech. Pub. No.128, 1952
- [4] Rush, A. I., Freeman, J. W., and White, A. E.: Abnormal Grain Growth in S-816 Alloy. NACA TN 2678, 1952.

# Thermodynamic Analysis of Cascade Refrigeration System using R508B - R404A for Vaccine Manufacturing Application

A. H. A. Shah, B. Dr. R. G. Kapadia

**Abstract**— COP of any refrigerating system decreases rapidly as the difference between the evaporating and condensing temperature increases. The single stage Vapour Compression Refrigeration (VCR) system is limited to an evaporator temperature of  $-40^{\circ}\text{C}$ . For an application which requires temperature below  $-40^{\circ}\text{C}$  than either cascade refrigeration system or multi stage VCR system is employed. For Vaccine Manufacturing Application the required temperature is in the range of  $-70^{\circ}\text{C}$  or below. This paper evaluates the thermodynamic analysis of cascade refrigeration system charged with ozone friendly refrigerants pair R508B-R404A and used for Vaccine Manufacturing Application. R508B is an azeotropic mixture composed of HFC refrigerants R-23/116 (46%/54% wt.). R404A is a ternary mixture composed of HFC refrigerants R-125/143a/134a (44%/52%/4% wt.). A parametric study is presented by varying Condensing, evaporating, subcooling and superheating temperatures in the HT and LT circuits and temperature difference in the cascade heat exchanger. Optimization of the overall system COP is evaluated with respect to intermediate temperature ( $T_{C,LT}$ ).

**Index Terms**—Cascade refrigeration system, COP, R404A, R508B.

## I. INTRODUCTION

Many industrial applications require low temperature cooling and high temperature heating simultaneously which cannot be achieved simultaneously and effectively by single stage or multistage systems due to individual limitation of refrigerant. Hence cascade system is the best alternative in these situations. A suitable selection of refrigerants used in HT and LT cycles can provide a large temperature lift while attaining good system efficiency. In low temperature refrigeration application including rapid freezing, storage of frozen food, liquefaction of petroleum vapour, manufacturing of dry ice etc. where the required evaporating temperature between  $-40^{\circ}\text{C}$  and  $-130^{\circ}\text{C}$ , the cascade system is often a convenient option, provided two suitable working fluids are chosen.

Vaccines are harmless agents, perceived as enemies. They are molecules, usually but not necessarily proteins, that elicit an immune response, thereby providing protective immunity against a potential pathogen. Production of Vaccine is achieved at the temperature of  $-70^{\circ}\text{C}$  or below [1]. To obtain this temperature cascade refrigeration systems are used. A commonly used refrigerants pair in the past has been R12, R502 in high temperature cycle and R13 in low temperature cycle of cascade refrigeration system. These refrigerants have been phased out since 1996 in the developed countries, and should be totally phased out by 2010 in developing countries as per Montreal Protocol [2] and its amendments from the United Nations Environment Programme (UNEP) [3]. It becomes impendence to look for new refrigerants to substitute for R12, R-502 and R13. Dinitrogen monoxide ( $\text{N}_2\text{O}$ ) was reported to be used as a cascade refrigerant for achieving temperatures around  $-80^{\circ}\text{C}$  [4]. Carbon dioxide ( $\text{CO}_2$ ) is also used in the low-temperature stage in two-stage cascade refrigeration systems [5], for which the lowest refrigeration temperature is limited above  $-55^{\circ}\text{C}$  because of its high triple-point temperature. Carbon dioxide and nitrous oxide binary mixture [6] is also used in low temperature cycle with R404a in high temperature cycle and experimental results were compared with R23 in low temperature cycle and R404a in high temperature cycle. Baolian Niu and Yufeng Zhang [7] used a new binary mixture of R744 and R290 as an alternative natural refrigerant to R13 and Experimental studies for this mixture on a cascade refrigeration system shows COP and refrigeration capacity of this binary mixture were higher than those of R13. Molenaar [8] compared the performance of three types cascade refrigerated test chambers. He concluded that R22 and R23 were acceptable substitutes for R502 and R13 in cascade refrigeration systems for test chamber application. Keumnam Cho, et al. [9] also investigated the effect of experimental parameters of a cascade system using R23 as alternative refrigerant. R508b is a binary azeotropic mixture comprised of R23 (46%) and R116 (54%) on a mass basis. It has many remarkable properties, such as non-flammable, non-toxic, etc. Di Nicola et al. [10] proposed blends of carbon dioxide and HFCs as working fluids for the low-temperature circuit in cascade refrigerating systems. Results showed that the R744 blends were an attractive option for the low-temperature circuit of cascade systems operating at temperatures approaching  $200^{\circ}\text{C}$ . K. Wadell [11] studied a two stage R134a/R508b cascade

- A. PG Student of Thermal Engineering at SVM Institute of Technology, Bharuch 392 001 India (e-mail: hirenshah29@hotmail.com)
- B. Professor in Mechanical Engineering Department at SVM Institute of Technology, Bharuch 392 001 India (e-mail: rag260475@gmail.com).

refrigeration system providing refrigerant (R508b) to the evaporator at flow rates of  $(0.833- 1.167) \times 10^{-3}$  kgs-1 and saturation temperatures of 187.15- 194.15 K. Di Nicola [12] has experimentally determined COP of carbon dioxide and nitrous oxide binary mixture in low temperature cycle with R404a in high temperature cycle and results were compared with R23 in low temperature cycle and R404a in high temperature cycle. The thermal design of condenser, cascade condenser and evaporator of cascade refrigeration system has been carried out using two HFC refrigerant pairs R404A-R508B and R410A-R23 by A. D. Parekh [13].

In this study, thermodynamic analysis of cascade refrigeration system using eco-friendly refrigerants pair R508B-R404A is carried out in terms of COP and Exergy. Independent variables have been varied to study their effect on performance parameters of the system with pre defined range of temperature which is shown in table I.

### II. CASCADE REFRIGERATION SYSTEM

Schematic diagram and P-h plot of two stage cascade refrigeration system are shown in Fig. 1 and Fig. 2 respectively.

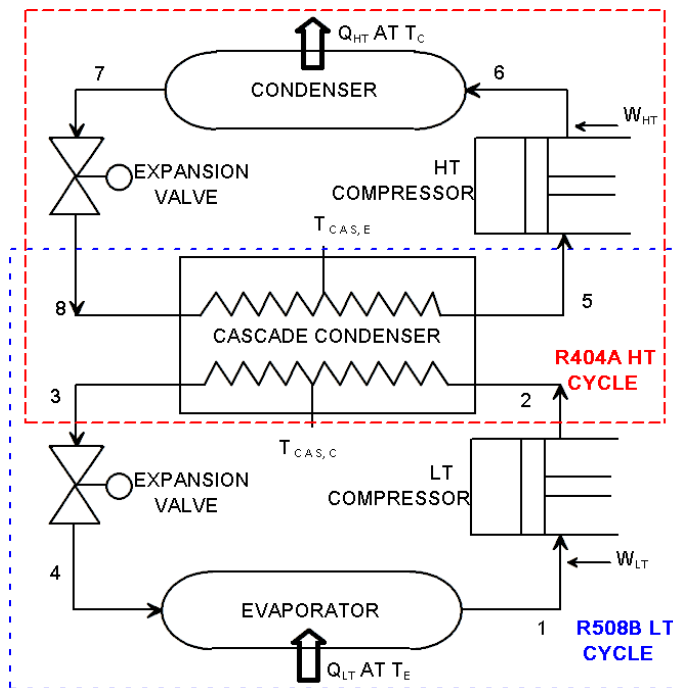


Fig. 1. Schematic diagram of a R508B-R404A cascade refrigeration system

The refrigerant pair R508B - R404A is used as LT - HT cycle. For the LT cycle, superheated refrigerant R508B vapour at state 1 is compressed to state 2 in the compressor and then cooled to saturated liquid at state 3 in the cascaded condenser. The liquid at state 3 is then expanded in an isenthalpic expansion device to state 4. The useful cooling in the evaporator is achieved by evaporating LT refrigerant R508B from state 4 to state 1. For the HT cycle, superheated refrigerant R404A vapour at state 5 is compressed to state 6;

superheated vapour is then cooled in the condenser to saturated liquid at state 7. The refrigerant R404A from state 7 is expanded to state 8 in an expansion device, followed by heating in the cascaded heat exchanger. The only useful refrigerating effect is produced in the evaporator of LT side of the cascade system. The HT and LT cycle used a refrigerant with high boiling temperature and low boiling temperature respectively.

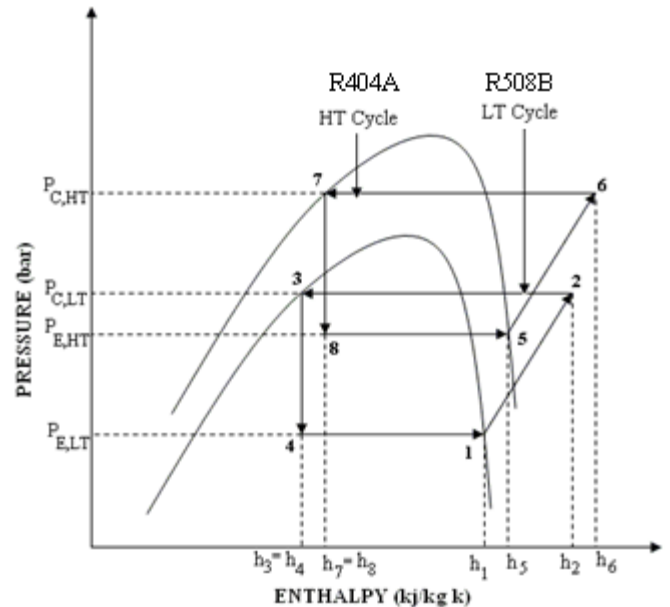


Fig. 2. Pressure-Enthalpy diagram of a R508B-R404A cascade refrigeration system

TABLE I  
TEMPERATURE RANGE OF PARAMETERS

Parameters	Temperature range ( $^{\circ}$ C)	
	Lower	Higher
Evaporator temperature of LT system ( $T_{E,LT}$ )	-90	-50
Condenser temperature of HT system ( $T_{C,HT}$ )	30	70
Cascade condenser temperature difference ( $\Delta T_{CC}$ )	0	10
Condenser temperature of LT system ( $T_{C,LT}$ )	-50	10
Degree of subcooling ( $\Delta T_{SUB}$ )	0	10
Degree of superheating ( $\Delta T_{SUP}$ )	0	10

### III. THERMODYNAMIC ANALYSIS

The thermodynamic analysis of R508B-R404A cascade refrigeration system was performed based on the following general assumptions:

- Refrigerants at the cascade heat exchanger outlet for HT cycle and evaporator outlet are saturated.
- Compression processes are adiabatic but not isentropic, with isentropic efficiency ( $\eta_{ISEN}$ ) of 0.8 for both LT and HT compressors.[13]
- Pressure losses in connecting pipes, compressor valves and heat exchangers are negligible.

- Heat transfer processes in all the heat exchangers are isobaric.
- Heat transfer with the surroundings is negligible for all the heat exchangers.
- Isenthalpic expansion of refrigerants in expansion valves.
- Changes in kinetic and potential energy are negligible.
- Cooling capacity of the system is 0.439 KW

Based on assumptions described above, the three balance equations are applied to find the work input to the compressor, the heat transfer rates of condenser; cascade condenser and the exergy destruction rate.

Mass balance 
$$\sum_{in} \dot{m} = \sum_{out} \dot{m} \tag{1}$$

Energy balance 
$$\dot{Q} - \dot{W} = \sum_{out} \dot{m} \cdot h - \sum_{in} \dot{m} \cdot h \tag{2}$$

Exergy balance 
$$\dot{X}_{des} = \sum_{out} \left(1 - \frac{T_0}{T_j}\right) \cdot \dot{Q}_j - \dot{W} + \sum_{in} \dot{m} \cdot \psi - \sum_{out} \dot{m} \cdot \psi \tag{3}$$

COP of the system can be defined as: 
$$COP = \frac{(h_5-h_3)(h_1-h_4)}{(h_5-h_2)(h_2-h_3)+(h_5-h_3)(h_2-h_1)} \tag{4}$$

The heat rejection ratio (HRR) can be expressed in terms of COP as: 
$$\xi = 1 + 1/COP \tag{5}$$

Hence, the refrigeration rate can be expressed in terms of HRR and efficiency as: 
$$R = \frac{\psi}{\eta(\xi-1)} \tag{6}$$

Additionally, the exergetic efficiency of the system yields: 
$$\eta_{II} = \frac{((T_0/T_{LT})-1)\dot{Q}_{LT} + (1-(T_0/T_{HT}))\dot{Q}_{HT}}{W_{LT} + W_{HT}} \tag{7}$$

COP, exergetic efficiency and exergy destruction for the cascade system has been evaluated from the above-mentioned thermodynamic model for R508B-R404A cascade refrigeration pair. A parametric analysis has been performed by varying subcooling and superheating, condensing temperature of LT cycle and isentropic efficiency of both cycles.

#### IV. RESULTS AND DISCUSSIONS

Computational Mathematical program has been developed to find the effect of particular parameter on the performance of a system by considering other parameters of a system as constant. EES [14] software package is used for the analysis of R508B-R404A cascade refrigeration system. The parameters have been varied for the computation results are mentioned in Table 1. The parameters taken for the analysis are given below.

- Evaporator temperature of LT system ( $T_{E,LT}$ ) = -85° C
- Condenser temperature of HT system ( $T_{C,HT}$ ) = 50° C
- Cascade condenser temperature difference ( $\Delta T_{CC}$ ) = 2° C
- Condenser temperature of LT system ( $T_{C,LT}$ ) = -27° C
- Degree of subcooling ( $\Delta T$ )<sub>SUB</sub> = 0° C in HT and LT cycle
- Degree of superheating ( $\Delta T$ )<sub>SUP</sub> = 0° C in HT and LT cycle

Fig. 3 shows the effect of LT cycle evaporator temperature ( $T_{E,LT}$ ) on COP; it is varied from -90° C to -50° C. All the other operating parameters are kept constant. As evaporating temperature is increased the pressure ratio is decreased so finally total work consumed by the compressor is reduced and refrigerating effect of the system is increased. COP of the system is also increased from 0.4334 to 0.7401.

Fig. 3 shows the effect of LT stage condensing temperature ( $T_{C,LT}$ ) on COP. COP of the system is initially increased to 0.6149 at -14° C after further increment in  $T_{C,LT}$  the COP is decreased. Work done and refrigerating effect both are decreased with increments of  $T_{C,LT}$ .

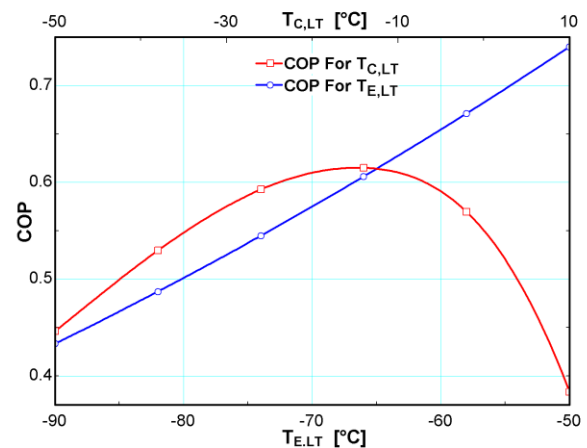


Fig. 3. Effect of LT cycle evaporating temperature and effect of LT cycle condensing temperature on COP of R508B-R404A cascade refrigeration system

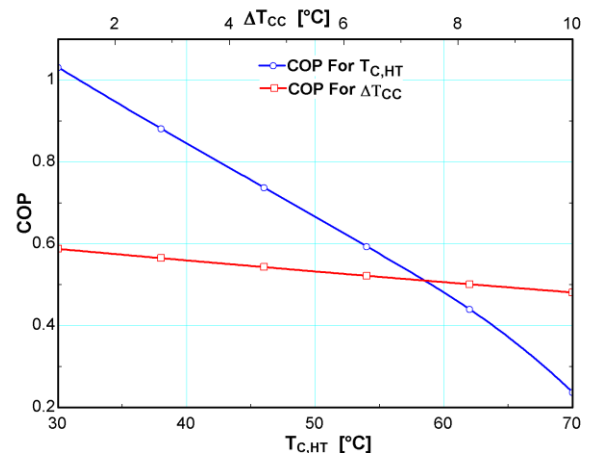


Fig. 4. Effect of HT cycle condensing temperature and effect of cascade condenser on COP of R508B-R404A cascade refrigeration system

Fig. 4 shows that the COP of the system is reduced from 1.031 to 0.2372 due to increased in HT cycle condenser temperature ( $T_{C,HT}$ ) from 30° C to 70° C. Compressor work is increased due to the increments of HT condensing temperature. Pressure ratio is also increased; so the COP of the system is decreased. The HT condensing temperature ( $T_{C,HT}$ )

$T_{HT}$ ) has no effect on the refrigeration effect therefore; it is remained constant.

Fig. 4 shows that when cascade condenser temperature ( $\Delta T_{CC}$ ) is increased from 1° C to 10° C; COP of the system is decreased about 18.10%. Work done of compressor is increased with  $\Delta T_{CC}$  but there is no change in refrigerating effect.

Fig. 3 and Fig. 4 depicts that the COP of R508B-R404A cascade system increased and decreased linearly with  $T_{E,LT}$  and  $\Delta T_{CC}$ ;  $T_{C,HT}$  respectively. COP of the system is increased up to certain point after that the COP is decreased. So, further analysis is carried out with the parameter  $T_{C,LT}$ .

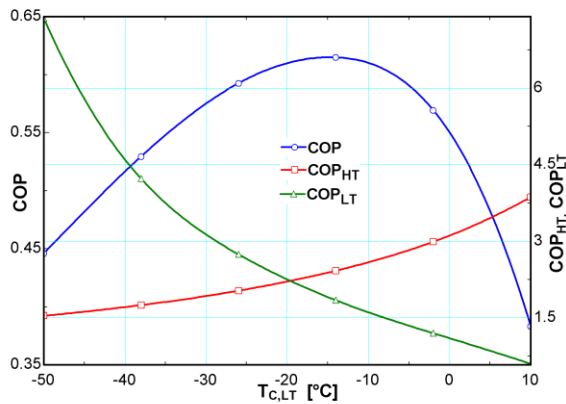


Fig. 5. Effect of LT cycle condensing temperature on COP, COP<sub>HT</sub> and COP<sub>LT</sub> of R508B-R404A cascade refrigeration system

Fig. 5 Shows the effect of LT stage condensing temperature ( $T_{C,LT}$ ) on COP, COP<sub>HT</sub> and COP<sub>LT</sub>. The COP of LT cycle is decreased with the increments of  $T_{C,LT}$  which is increased from -50° C to 10° C; while COP of HT is increased with  $T_{C,LT}$ .

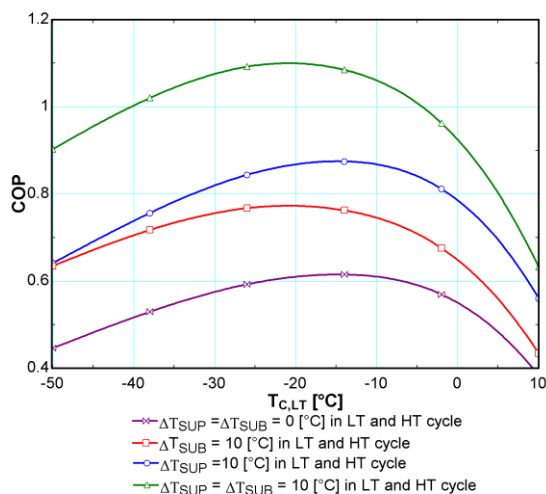


Fig. 6. Effect of LT cycle condensing temperature on COP of R508B-R404A system considering at 0° C to 10° C subcooling and superheating in both cycles

Fig. 6 shows the variation in COP of system at different value of superheating and subcooling at the condensing temperature of LT cycle is varied from -50° C to 10° C. Maximum COP of the system is increased slightly with an increased in subcooling while increased significantly with increased in superheating. For zero degree of subcooling and superheating COP<sub>MAX</sub> is 0.575. It is increased to 0.8748 at 10 degree of superheating and also increased to 0.7673 at 10 degree of subcooling in both cycle. It is again increased to 1.092 at 10 degree of subcooling and superheating in both cycles.

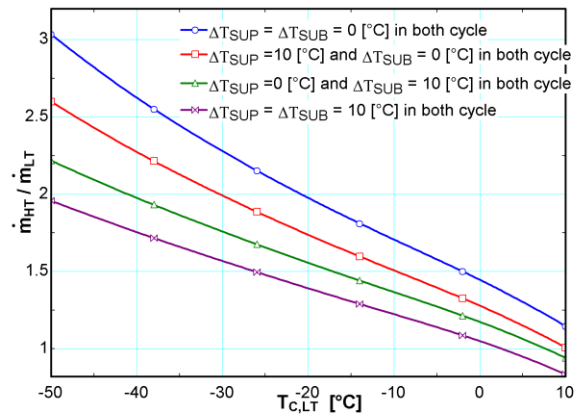
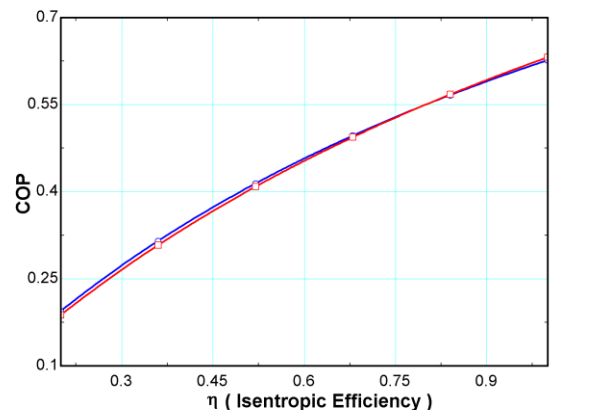


Fig. 7. Effect of LT cycle condensing temperature on mass flow ratio of R508B-R404A cascade refrigeration system varying at 0 to 10 degree of superheating and subcooling

Fig. 7 shows that when condensing temperature of LT is increased from -50° C to 10° C; the mass flow ratio ( $\dot{m}_H/\dot{m}_L$ ) is decreased. At higher  $T_{C,LT}$ , the difference in values of mass flow rate reduces for different degree of subcooling and superheating. Ratio of mass flow ( $\dot{m}_H/\dot{m}_L$ ) is 3.035 in case of 0 degree of subcooling and superheating in both cycles. It is reduced to 2.599, 2.217 and 1.958 in case of 10° C superheating, 10° C subcooling and 10° C subcooling-superheating in both HT and LT cycles respectively.



— R508B ( $T_{C,LT} = -21^\circ\text{C}$ ,  $T_{E,LT} = -84^\circ\text{C}$ ,  $\Delta T_{SUB} = \Delta T_{SUP} = 0^\circ\text{C}$ ,  $\eta_{R404A} = 0.8$ )  
 — R404A ( $T_{C,HT} = 50^\circ\text{C}$ ,  $T_{E,HT} = -23^\circ\text{C}$ ,  $\Delta T_{SUB} = \Delta T_{SUP} = 0^\circ\text{C}$ ,  $\eta_{R508B} = 0.8$ )

Fig. 8. Effect of having different isentropic efficiency in the compressor of both cycles.

Fig. 8 shows the effect of various isentropic efficiency in R508B cycle and R404A cycle with respect to COP. Graph shows that an increase in isentropic efficiency of R508B results in a significant increase in system COP from 0.1945 to 0.6267, which is 68.96% considering other parameters constant. In the same way an increase in isentropic efficiency of R404A results in increase in system COP from 0.1878 to 0.6316, which is 70.27% considering other parameters constant. This leads to a conclusion that considering same isentropic efficiency for both the cycle compressors are practical and make simpler the calculation of system COP.

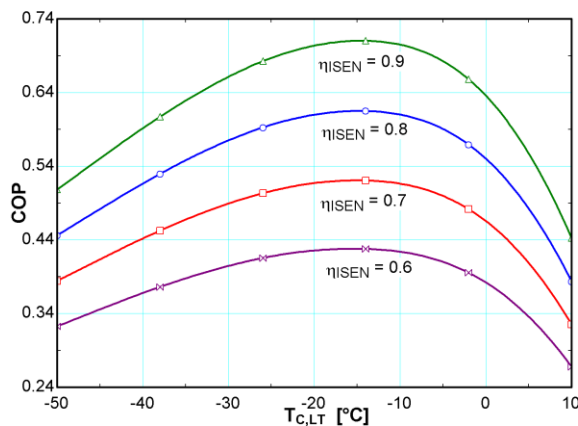


Fig. 9. Effect of LT stage condensing temperature on COP at different isentropic efficiency of both HT and LT stages compressors for R508B-R404A system

Fig. 9 shows the variation in COP with low stage condensing temperature ( $T_{C,LT}$ ) for different isentropic efficiency (60%, 70%, 80% and 90%) of the system for both HT and LT cycle compressors. From graph maximum COP of the system goes upward proportionally with an increase in isentropic efficiency ( $\eta_{ISEN}$ ).

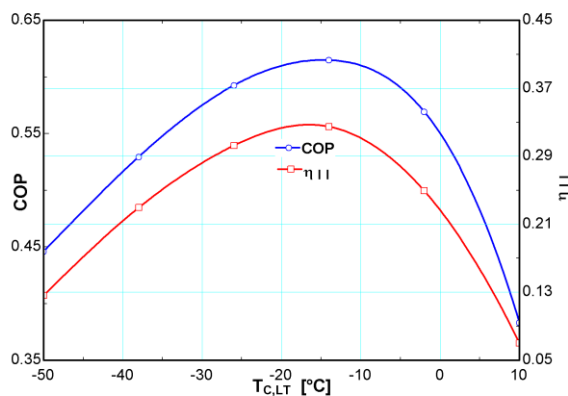


Fig. 10. Effect of LT stage condensing temperature on COP and exergetic efficiency of both HT and LT stages compressors for R508B-R404A system

From fig. 10, it can be seen by increasing the condensing temperature in LT stage, due to the increased in refrigeration effect; COP and exergetic efficiency ( $\eta_{II}$ ) of the cascade system increased up to certain value of  $T_{C,LT}$ . The exergetic efficiency and COP of the system may be maximum at a certain value of  $T_{C,LT}$ , as shown in Fig. 10, in which an  $T_{C,LT}$  optimum, exists is  $-14^{\circ}C$ .

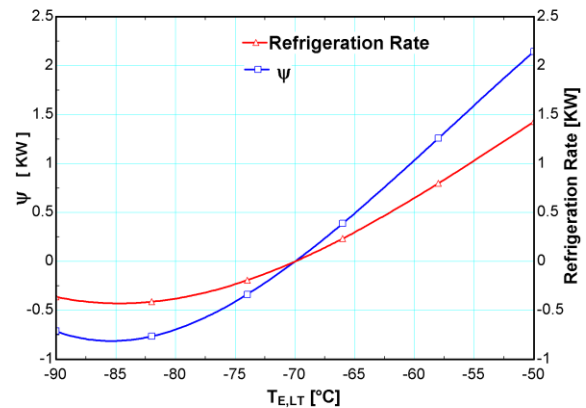


Fig. 11. Effect of LT stage evaporating temperature on Exergy of the system and refrigeration rate for R508B-R404A system

Fig. 11 shows the effect of  $T_{E,LT}$  on the exergy rate and refrigeration rate of the system. The exergy rate of the system is minimum at a certain  $T_{E,LT}$ , as shown in Fig. 11, in which an optimal  $T_{E,LT}$ ; exists is  $-70^{\circ}C$ . At the temperature of  $-70^{\circ}C$  ( $T_{E,LT}$ ) the exergy rate and refrigeration rate is the same.

It is interesting to note that at this point if HRR of the system, the ambient temperature and the reservoir temperatures are kept constant (efficiency is constant), the rate of exergy output will be directly proportional to refrigeration rate. This means that both exergy rate and refrigeration rate reach maxima at the same point and hence maximizing either would maximize the other. Fig. 11 shows the trend of refrigeration rate with varying  $T_{E,LT}$  for the given operating parameters. The trend is exactly similar to that for exergy rate of the system.

V. CONCLUSION

This paper presented a thermodynamic analysis of R508B-R404A cascade refrigeration system for vaccine manufacturing application. Thermodynamics analysis has been presented the following translation.

- The COP of the R508B-R404A system decreased from 1.031 to 0.2372 at HT cycle condenser temperature ( $T_{C,HT}$ ) is varied from  $30^{\circ}C$  to  $70^{\circ}C$  while other parameters remained constant.
- The COP of R508B-R404A system reduced by about 18.11% when temperature difference in cascade condenser ( $\Delta T_{CC}$ ) was increased from  $1^{\circ}C$  to  $10^{\circ}C$ .
- The COP of the R508B-R404A system increased from 0.4334 to 0.7401 at LT cycle evaporator temperature ( $T_{E,LT}$ ) is varied from  $-90^{\circ}C$  to  $-50^{\circ}C$  while other parameters remained constant.

- The COP of HT cycle increased whereas the COP of LT cycle decreased when  $T_{C,LT}$  is varied from  $-50^{\circ}\text{C}$  to  $10^{\circ}\text{C}$  for R508B-R404A system. Overall COP of the R508B-R404A system is increased parabolic with the increment in LT stage condensing temperature ( $T_{C,LT}$ ) from  $-50^{\circ}\text{C}$  to  $10^{\circ}\text{C}$ . Maximum value of  $\text{COP}_{\text{MAX}}$  is 0.6149 at the optimum condensing temperature ( $T_{C,LT}$ ) is  $-14^{\circ}\text{C}$ .
- Maximum COP of the system R508B-R404A increased significantly with an increase in degree of superheat while increased slightly with increase in subcooling.
- An increase in condensing temperature ( $T_{C,LT}$ ) from  $-50^{\circ}\text{C}$  to  $10^{\circ}\text{C}$ ; mass flow ratio ( $\dot{m}_H / \dot{m}_L$ ) is decreased for different degree of superheating and subcooling for R508B-R404A system.
- An increase in isentropic efficiency ( $\eta_{\text{ISEN}}$ ) of compressors increased COP in parabolic way for R508B-R404A system.
- As  $T_{C,LT}$  is increased; COP and exergetic efficiency of the system reached at maximum value for an optimum value of  $T_{C,LT}$  as  $-14^{\circ}\text{C}$ .
- Exergy destruction rate of the system was minimum 0.08 KW at optimum  $T_{C,LT}$  is  $-14^{\circ}\text{C}$ .

## NOMENCLATURE

h	enthalpy (kJ/kg)
$\dot{m}$	mass flow rate (kg/min)
ODP	ozone depletion potential
$\dot{Q}$	heat transfer rate (KJ)
R	refrigeration rate
S	entropy (kJ/kg K)
T	temperature ( $^{\circ}\text{C}$ )
$\dot{W}$	rate of power consumed (KJ/s)

## Acronyms

CFC	chloro-fluorocarbon
COP	coefficient of performance
GWP	global warming potential
HFC	hydro-fluorocarbon

## Greek Symbols

$\eta$	efficiency
$\eta_{\text{II}}$	exergetic efficiency
$\dot{X}_{\text{des}}$	exergy destruction (KW)
$\xi$	HRR
$\psi$	rate of exergy output (KW)

## Subscripts

C	condenser
CC	cascade condenser
CAS	cascade heat exchanger
E	evaporator
HT	high temperature
ISEN	isentropic
LT	low temperature
MAX	maximum
SUB	subcooling
SUP	superheating

SUP superheating References

## REFERENCES

- [1] Roger Anderson, Anik Egan Gillian Chaloner, "A WHO guide to good manufacturing practice (GMP) requirements," World Health Organization, Geneva, Health, 1997.
- [2] Montreal protocol on substances that deplete the ozone layer, United Nations Environment programme (UNEP), 1987.
- [3] UNEP Assessment report of the technology and economic assessment panel, UNEP Ozone Secretariat, Nairobi, Kenya 2007.
- [4] Kruse,H.,Ru' ssmann,H., "The natural fluid nitrous oxide – an option as substitute for low temperature synthetic refrigerants", *International Journal of Refrigeration*, 29 (5), 2006, pp. 799–806
- [5] Nicola G.D., Giuliani G., Polonara F., Stryjek, R., "Blends of carbon dioxide and HFCs as working fluids for the low temperature circuit in cascade refrigerating systems", *International Journal of Refrigeration*, vol.28, 2005, pp. 130–140
- [6] DiNicola, G., Giuliani, G., Polonara, F., Santori, G., Stryjek, R., "Cascade Cycles Operating with CO2 + N2O Binary Systems as Low Temperature Fluid: Experimental Results", *International Conference of Refrigeration*, Beijing 2006.
- [7] Baolian Niu, Yufeng Zhang, "Experimental study of the refrigeration cycle performance for the R744/R290 mixtures", *International Journal of Refrigeration*, vol. 30, 2007, pp 37-42.
- [8] George L. Molenaar, "Use of R-22/23 in lieu of R502/13 in a cascade refrigeration system", in: *Proceedings of ASMENTC01 35th National Heat Transfer Conference*, Anaheim, California, 2001, pp.243-246.
- [9] Keumnam Cho, Jonghoon Park, Honggi Cho., "Performance of the cascade system using alternative refrigerants", in: *Proceeding of ASMENTC01 35th National Heat Transfer Conference*, Anaheim, California, 2001, pp. 595-P600,
- [10] Giovanni Di Nicola, Giuliano Giuliani, Fabio Polonara, Roman Stryjek., "Blends of carbon dioxide and HFCs as working fluids for the low temperature circuit in cascade refrigerating systems", *Int J Refrigeration*, Vol 28, 2005, pp. 130-140.
- [11] Robert Paul Wadell., Design of compact evaporators for ultra low temperature thermal management of microprocessors, Master thesis, Georgia Institute of Technology, 2005.
- [12] DiNicola, G., Giuliani, G., Polonara, F., Santori, G., Stryjek, R., "Cascade Cycles Operating with CO2 -N2O Binary Systems as Low Temperature Fluid: Experimental Results". *International Conference of Refrigeration*, Beijing 2007.
- [13] A.D.Parekh, P.R.Tailor, "Thermodynamic Analysis of R507A-R23 Cascade Refrigeration System", in *Proceeding of International Journal of Aerospace and Mechanical Engineering*, 2012, 6:1.
- [14] EES: Engineering Equation Solver 2006.

# Introduction and Mathematical modeling of Wall Climbing Robot

A. Chirag R Vyas, B. Vishal N Vyas

**Abstract**— This paper presents review of different wall climbing robots. Brief overview of their mechanism, advantages and disadvantages are discussed. A various aspects in mechanical design including locomotion mechanism, adhesion mechanism are discussed. Steps to derive generalized equation to find velocity and acceleration of wall climbing robot are discussed.

## I. INTRODUCTION

In recent time, there have been strong demands that use of robots for defense, surveillance, and counter terrorism missions. Most of the mobile robots now days essentially move in 2D plane without wall climbing capability. Wall climbing robot is one of the special in mobile robots because it moves in vertical plane.

Autonomy of a mobile robot means in practical terms that the robot acts autonomously means it navigates and fulfill Today ,in modern cities there is a great amount of different and complicated shapes in high rise building. However ,these external cladding walls require constant cleaning. As a result, even skilled workers with safety ropes have difficulties in climbing those buildings and currently almost of them are still cleaned manually

. The development of climbing robots offers a novel solution to the abovementioned problems. To increase the efficiency in dangerous environments or difficult to access places, and to protect the human health and safety in hazardous tasks at that time we can use wall climbing robot like for inspecting nuclear power plants, outer surface of air planes, pipes, buildings, and leakages of gas tanks, and cleaning .it is also used this robot for search and rescue, surveillance, and entertainment

In this work, discussed about different challenges faced to make wall climbing robot. Review of different mechanism used to make wall climbing robot. Mathematical model for two differential drive body is moving in vertical surfaces.

## II. VARIOUS WALL CLIMBING ROBOTS

To make wall climbing robot, choosing right adhesion system is most important. So far many of researchers developed this type of adhesion technique, Magnetic Adhesion, Elastomer Adhesion, suction-pad adhesion.

W.Fischer, F. Tache and R. Siegwart [1] describes a novel solution to a mobile climbing robot on magnetic wheels, designed for inspecting the interior surfaces in gas tanks made out of thin metal sheets. The simplest structure for a climbing robot is a vehicle on two or more magnetic wheels. These wheels increase the normal force to the ground and thus allow for being more independent from the direction of gravity .A cylindrical magnet in the middle; two plates out of magnetic steel and of slightly bigger diameter on both sides (to better conduct the magnetic force into the surface) and a thin layer of rubber around the steel plates (to increase the friction to the ground from ( $\mu=0.3-0.5$  to  $\mu =0.5-0.8$ ). This robot used active adhesion system and wheeled locomotion system. This robot has capacity to work all inclinations of the surfaces and only used for thin metal surfaces.

A wall climbing robot with vacuum caterpillar wheel system operated by mechanical valve [2], this robot consists of vacuum caterpillar wheel and main body including vacuum pump and power supply. Especially, vacuum caterpillar wheel is composed of caterpillar wheel on which 24-suction pads are installed, mechanical valves opening and closing vacuum air flow, grooved guide-rail which guides the movement of suction pads according to the rotation of the wheel and controls the operation of mechanical valves, and rotary joint preventing pneumatic tubes from twisting by wheel rotation.

O. Unver and M. Sitti [4] describes energy efficient, light weight and robust climbing robot .It has a continuous attachment and detachment method for stable climbing on many surfaces using a at bulk, tacky elastomer adhesive. The main advantages of this attachment method over the others are its energy efficiency, quiet operation, and surface independent adhesion if the roughness is on the order of a micron or smaller scale. Elastomer adhesives conform to a given surface roughness when pressed vertically. In the contact region, intermolecular forces, such as van-der-Waals and hydrogen bonding forces , give high attachment forces. Therefore, these elastomer can work on a wide range of surface materials and environments. Once the adhesive material contacts a given surface, it can stay attached without

A. Mechanical Engineering Department, Gandhinagar Institute of Technology, Khatraj, Gandhinagar, Gujarat, 382721, India (Email: [vyaschirag88@gmail.com](mailto:vyaschirag88@gmail.com))

B. Mechanical Engineering Department, Gandhinagar Institute of Technology, Khatraj, Gandhinagar, Gujarat, 382721, India, (Email: [vishuvyas2003@gmail.com](mailto:vishuvyas2003@gmail.com))



any power consumption. For power efficient detachment, the elastomer should be peeled by curling back from the edge

This paper[5] presents a new taxonomy for wall climbing robots, among which L-NP and NPT methods with continuous and rapid locomotion are studied emphatically.

In wall climbing robots domain, adhesion principle and locomotion determine many performance indexes such as reliability, maneuverability, payload capacity, endurance distance and so on. Pneumatic adhesion principle uses a prime motor to drive power machines to form inerratic air attribute (velocity, pressure) transformations, which sequentially produces a steady force pressing robot to the wall.

Negative pressure method uses fluid machinery such as vacuum motor, centrifugal impeller to suck the air from the inside the chamber and create vacuum inside the chamber, there are two types of its single cup approach and multi cup approach.

### III. CHALLENGES TO MAKE WALL CLIMBING ROBOT

In climbing robots it is real challenge to carry its own weight because the work is done in opposite direction of gravity. The most challenging tasks in climbing robot design is to develop proper adhesion mechanism to ensure that the robot sticks to wall surfaces reliably without sacrificing mobility. Locomotion system selection is depends upon selection of adhesion system.

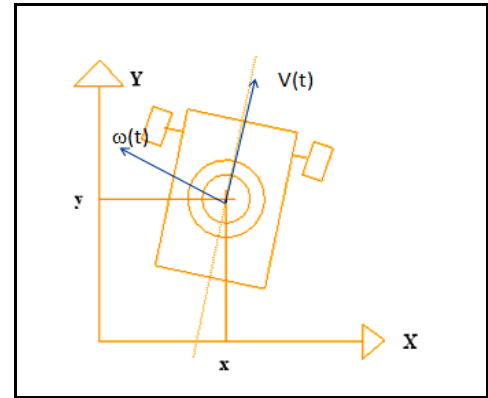
Robot is moving on the wall or it is in steady condition it has enough adhesion force to stick with the wall without affecting its mobility. Wall climbing robot should be in light weight, compact, safety and inexpensive.

### IV MATHEMATICAL MODELING

For two wheeled robot design [3] the derivation of equation for position, velocity and angular acceleration for vertical surfaces. Inorder to achieve pure rolling action of wheels(without sliding and slipping), torque is required on each wheel so the power supply to each wheel is given with different motors.

STEPS:-

- Start to derive the model for robot chassis
- Derive kinematic and dynamic equations for robot chassis
- Apply Euler-Lagrange formulation for robot chassis
- Equations for wheel dynamics
- Formulate combined chassis and wheel model equation
- Total energy of the system
- Apply Euler-Lagrange formulation for the whole system for the dynamics.



Where,  $x(t)$ ,  $y(t)$  are robot position in  $x$  and  $y$  coordinates and  $v(t)$ ,  $\omega(t)$  Robot linear and angular velocity w.r.t body fixed frame. Kinetic energy(K.E) of the robot chassis is the sum of linear motion and angular motion. Consider potential energy(P.E) in terms of height( $y$ ) at time  $t$ . Total energy(T.E) of the robot chassis is,

$$T.E=K.E+P.E$$

$$\frac{1}{2}McV^2 + \frac{1}{2}Ic\omega^2 + \frac{1}{2}MwV(l)^2 + \frac{1}{2}MwV(r)^2 + \frac{1}{2}Iw\omega(l)^2 + \frac{1}{2}Iw\omega(r)^2 \tag{1}$$

Put the relationship value of the velocities,

$$\left( \frac{1}{2}Ic \frac{Rw^2}{Rc^2} + \frac{1}{8}McRw^2 + \frac{1}{2}MwRw^2 + \frac{1}{2}Iw \right) \omega(l)^2 + \left( \frac{1}{2}Ic \frac{Rw^2}{Rc^2} + \frac{1}{8}McRw^2 + \frac{1}{2}MwRw^2 + \frac{1}{2}Iw \right) \omega(r)^2 + \left( \frac{1}{4}McRw^2 - Ic \frac{Rw^2}{Rc^2} \right) \omega(l)\omega(r) \tag{2}$$

Where

$Mc$  is robot chassis mass,

$Mw$  wheel mass,  $Rc$  is chassis radius,

$Rw$  is wheel radius,  $Iw$  wheel moment of inertia,

$Ic$  chassis moment of inertia.

Potential energy of the robot body is  $Mgy$ .

Apply Euler-Lagrange formulation, total energy becomes

$$\begin{aligned} & \left[ I_c \frac{Rw^2}{Rc^2} + \frac{1}{4} McRw^2 + MwRw^2 + Iw \right] \dot{\omega}(l) + \\ & \left( \frac{1}{4} McRw^2 - I_c \frac{Rw^2}{Rc^2} \right) \dot{\omega}(r) - \tau(l) - \frac{1}{2} Mgy \\ & \left[ I_c \frac{Rw^2}{Rc^2} + \frac{1}{4} McRw^2 + MwRw^2 + Iw \right] \dot{\omega}(r) + \\ & \left( \frac{1}{4} McRw^2 - I_c \frac{Rw^2}{Rc^2} \right) \dot{\omega}(l) - \tau(r) - \frac{1}{2} Mgy = [0,0] \end{aligned} \tag{3}$$

Where

$\tau(l)$ ,  $\tau(r)$  torque for left and right wheel

Left wheel acceleration equation is,

$$\begin{aligned} & [(4IcRw^2 + McRw^2 Rc^2 + 4MwRw^2 Rc^2 + 4IwRc^2) \tau(l) \\ & + (-McRw^2 Rc^2 + 4IcRw^2) \tau(r) + \\ & (4IcRw^2 + 2MwRw^2 Rc^2 + 2IwRc^2) Mgy] \\ & \frac{[(4Mw^2Rw^4 Rc^2 + 4Iw^2 Rc^2 + 4MclcRw^4 + \\ & 8MwIcRw^4 + 8IcIwRw^2 + 2McMwRw^4 Rc^2 \\ & + 2McIwRw^2 Rc^2 + 8MwIwRw^2 Rc^2)]}{\phantom{}} \end{aligned} \tag{4}$$

Similarly can be derived for  $\dot{\omega}(r)$

$$\text{Linear acceleration, } \dot{v}(t) = \frac{1}{2} Rw \dot{\omega}(l) + \frac{1}{2} Rw \dot{\omega}(r) \tag{5}$$

$$\text{Angular acceleration, } \dot{\omega}(t) = \frac{Rw \dot{\omega}(l) - Rw \dot{\omega}(r)}{Rc} \tag{6}$$

These equations are useful to find velocity and acceleration for two wheeled vertical moving vehicle.

### V. CONCLUSION

The literature is reviewed for various adhesion system and locomotion mechanism and concluded that single cup approach is best suitable for different wall climbing surfaces as well as in mobility. Derived mathematical model for two differential wheeled drive robot is moving in vertical surfaces.

### REFERENCES

- [1] F Tache W.Fisher and R.Siegrwart. Magnetic wall climbing robot thin surfaces with specific obstacles. 6th international conference on field and service robotics, Chamonix: France, 2007.
- [2] Dongmok K. Hojoon Y. Kyouhee L. Kunchan S. Doyoung C. Jongwon K. Hwan, K. A wall climbing robot with vacuum caterpillar wheel system operated by mechanical valve. 9th international conference on climbing and walking robots, 2006
- [3] Maple soft a division of Waterloo Maple Inc. Mobile robot modeling and simulation
- [4] O. Unver and M. Sitti. Tankbot:a miniature,a peeling based climber on rough and smooth surfaces.
- [5] Z. Jiang, J. Li, X. Gao, N. Fan, and Wei B. Study on pneumatic wall climbing robot adhesion principle and suction control. International Conference on Robotics and Biomimetics Bangkok, Thailand, pages {1812-1817}, 2009

# Experimental Investigation of material of ball valve by using FEA analysis

A. G. N. Parmar, B. S. P. Patel, c. A. A. Darji

**Abstract**— In this paper, we use the Pro engineer model of ball valve. The objective of this project is to study design of ball valve and increase life of ball valve and also optimization of material . For that we used different standards like ASME (American standard of mechanical engineer), API (American petroleum institute)etc. in this project we check the design of specification of valve, then we check its wall thickness. Generally most of valve has high wall thickness. In this project we recommended new material for particular design and we reduce valve weight. For that selection we use finite element analysis of ball valve and after analysis of valve we get our main objective. In this analysis we use proe-engineer software for modeling purpose and we use ANSYS software for analysis purpose.

**Index Terms**—Material, Product Design, FEA, Ball Valve

## I. INTRODUCTION

Products designed for recognized applications with problem-solving design, expected quality, believable life is of high demand and has become a mortal of human living purchasing a material with a restricted chemical composition and defined mechanical properties and utilizing at conceptual stage of product development is the present scenario. For the selected material, the required shape (casting, forging, plate, and pipe) and the possible manufacturing processes should be considered first. Then in accordance with the design code (e.g., ASTM, ANSI BIS) adopted, the specification for the material to be applied is selected and the allowable design stresses at the operation temperature, are determined. In correlation with design procedure, the proved uneconomical process in terms of cost and time namely physical model or prototype forms the basis for better understanding of a particular process. Furthermore, it is very difficult to develop prototypes for complex and multi dimensional products and then subjecting the same for property and functionality testing. A designed component is a unique part or subassembly that must be specifically designed and fabricated as part of the design being developed. Choosing between designed and standard components can be an important consideration. The choice between designing a special purpose component, optimized for performance and weight,

or purchasing a standard, “off-the-shelf” component from a supplier, can have far reaching performance, cost, quality and timing consequences. Market availability is lofty for multiple design build test cycles involved in the creation of a design. The product life depends highly upon the type and property of the material and in turn the properties are forbidden for the effective functionality It is obvious that chemistry of the alloy plays a vital role in determination of material parameters. As of functionality of the product is concerned, material characterization attributes involves the complex interrelation of a number of variables associated with product design, its manufacturing process and service conditions. In cases where design is typically too risky to develop , a first round of planning, material selection, functional design manufacture and assembly is undertaken to produce a model, which when tested may indicate ways of improving the product design. The current study is focused on imparting this information in conjunction with possibly extensive computer aided analysis and optimization either to manufacture an improved model for tests or to manufacture a prototype or the product itself. On thorough review over various literatures it is proved to be worthwhile to blend relationships between the domains of material design and product design in spite of various other considerable governing factors of design.

## II. SCOPE OF RESEARCH

### A. Model Geometry of ball valve

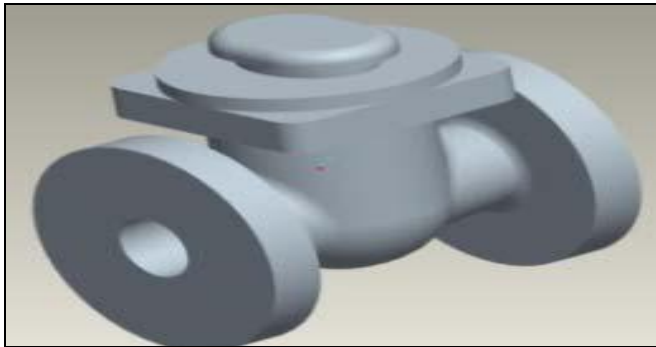
For this purpose artificial neural network has evolved as an efficient tool for developing models which costs less compared to traditional process [15]. To generate a new material with multiplicity in composition is too tedious and this proves to be a barrier for the development of many new products. As a remedy for this deficient and headed for the formulation of innovative materials, survival of ground-breaking procedures is a must. The properties used are those involved in the basic formulas of mechanical design and those typically used to characterize material properties provided in engineering handbooks specifications and supplier catalogs. Designers are typically interested in strength, ductility, toughness Young’s modulus and fatigue strength. These models states that the combination of certain properties may impose performance limitation s in addition to those constraints imposed by individual properties The scope of this research on artificial intelligence based modeling technique is to optimize the ASTM A216 Gr. WCBhigh pressure application material with necessitated improvements in

- A. M.E Student, Mechanical Engineering Department, L.D.R.P- I.T.R, Gandhinagar, (e-mail: [gaurang.346@gmail.com](mailto:gaurang.346@gmail.com)).
- B. Asst. Prof. Mechanical engineering department, L.D.R.P-I.T.R, Gandhinagar, (e-mail: [satyampatel@yahoo.co.in](mailto:satyampatel@yahoo.co.in)).
- C. Asst. Prof., mechanical Engineering department, L.D.R.P-I.T.R, (e-mail: [ankitdarjildrp@gmail.com](mailto:ankitdarjildrp@gmail.com))

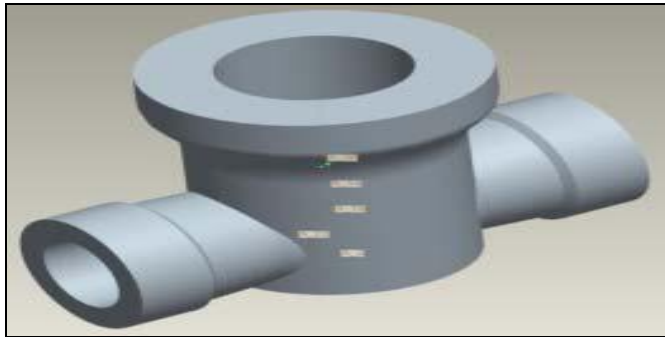
properties and chemical mixture of the existing alloys as per the functional requirements of the valves as shown in figure 1(a) and (b). solid model geometry for Gate & Ball Valve

Figure. 1(a)

Figure. 1(b)



B. Modeling Technique



In this paper first we get the product design of ball valve. Then we have to do modeling in pro-engineer software. for that we make all model of part like body, stem ,sheet ,stem ring, ball, nut part, than we make its assembly of all this part. Then after we do the analysis on this assembly model in ANSYS software. In ANSYS software we make structure analysis.

III. EXPERIMENTAL SCHEME

Experiments are conducted in a foundry for the ASTM A487 Gr 4C: 0.2 to 0.3 % Carbon; 0.4 to 0.8 % Silicon; 0.8 to 1.0 % Manganese; 0 to 0.03 % Sulphur; 0 to 0.03 % Phosphorus; 0.4 to 0.8 % Chromium; 0.4 to 0.8 % Nickel 0.15 to 0.3 % Molybdenum; maximum 0.5% Copper; 0 to 0.03 % Vanadium; 0 to 0.1 % Tungsten and a maximum of 0.60% of unspecified alloying elements. The melting range of the specified steel is about 2740-28000F. The test castings were poured in CO2 – sodium silicide moulds knocked out, risers were removed and then subjected to normalising, tempering sort of heat treatments. A sample of the data set is made available in table 1.

Table1: Set of sample data used for model development

Sl No	Chemical Composition (in %)											Mechanical Properties					
	C	Si	Mn	S	P	Cr	Ni	Mo	Cu	V	W	TS	YS	EI	RA	IS	Test temp °C
1	0.26	0.51	0.81	0.011	0.014	0.54	0.54	0.24	0.03	0.002	0.005	639	489	23.0	45.8	42	-46
2	0.23	0.55	0.84	0.010	0.015	0.50	0.45	0.18	0.03	0.003	0.008	635	509	22.6	36.0	58	-46
3	0.24	0.63	0.91	0.021	0.022	0.51	0.46	0.19	0.04	0.003	0.010	646	451	20.0	45.0	34	-46

Behavior of material depends on percentage composition of alloying elements and so the same is projected out for 250 tests wherein odd experimental observations with missing or contradictory interpretations and of abnormal strength values were omitted from the training data set. The material attributes like Yield Strength (YS) in Mpa Tensile Strength (TS) in Mpa, Percentage Elongation (EI), Percentage Reduction in Area (RA), and Impact Strength (IS) of the specimens were experimented. In routine utmost care is taken to maintain the test temp at -46°C (for impact testing). Equipped with the CHARPY Impact testing machine for test specimens prepared as per the standard (ISO CVN 2 mm SIZE - 10 x 10 x 55 mm) the impact tests are accomplished.

IV. FINITE ELEMENT ANALYSIS

As future prospects, the ability of the proposed valve material in relation to fatigue characteristics is identified using Finite Element Analysis by applying the observed results as input parameters in adoption to variation in thickness of ball valve and outer diameter of gate valve. The same is carried over to assess the stress concentration and to verify the contact status between the ball and the valve body finite element method is used Assessments of the stress concentration, strain distribution. Total deformation acting on the ball and gate valves for its different valve position is carried out. The output results for both valve types credited through finite element analysis are illustrated in figures.

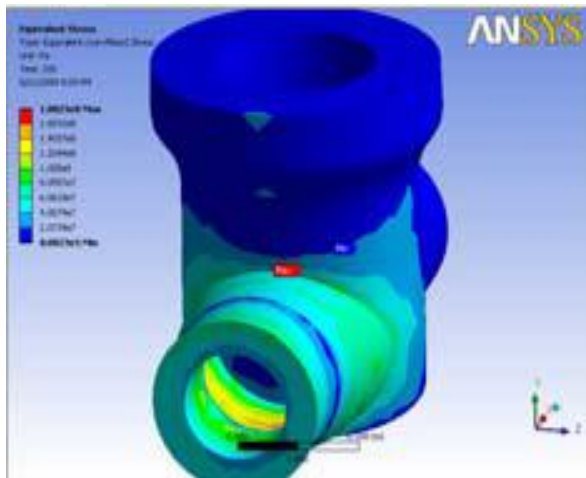


Figure 4 (a)

Fig 4 (a) & 4(b) shows sample analysis results for Ball Valve of 21 mm thickness 30 & 45 degree opening modes respectively

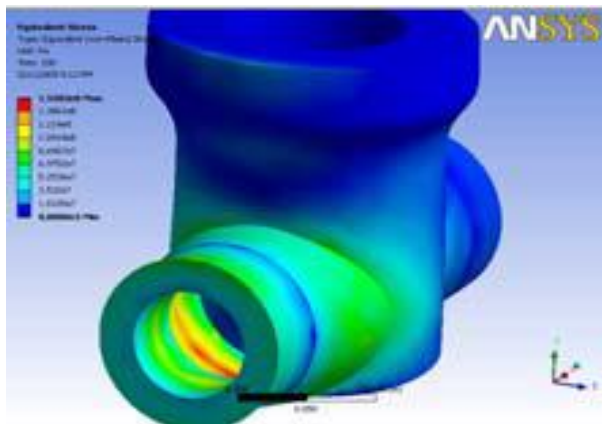


Figure 4 (b)

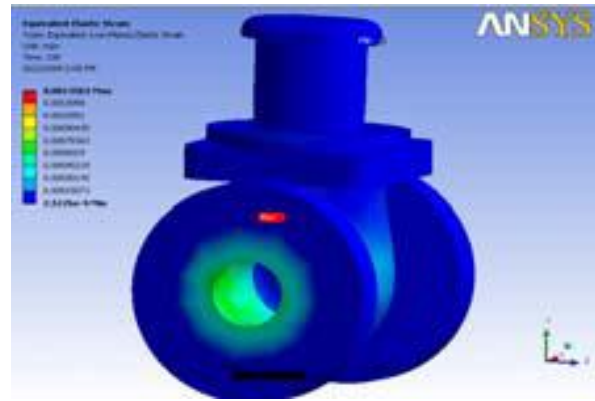


Figure. 5(b)

Fig 5 (a) & 5 (b) shows sample analysis results for valve of 90 mm OD for 10 mm and 70 mm openings respectively

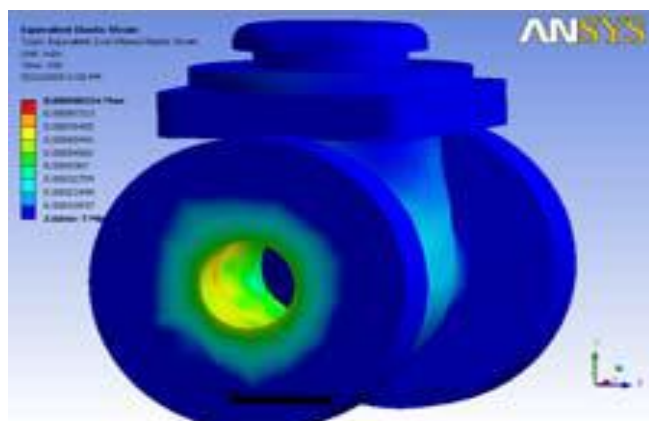


Figure. 5 (a)

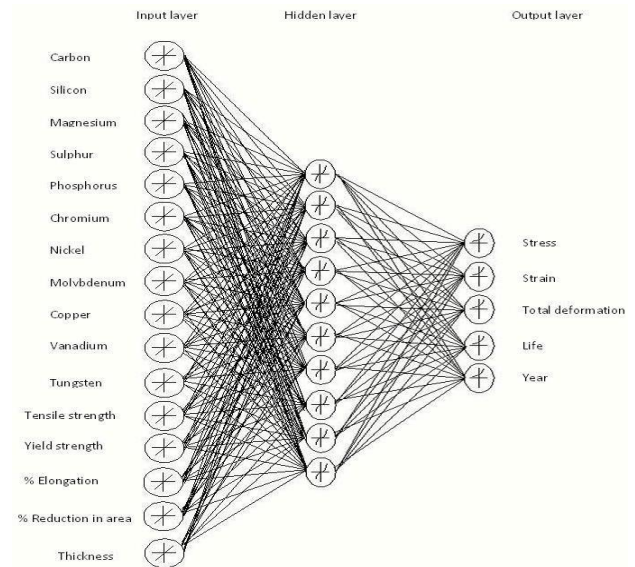


Figure. 6 Illustrates typical network structure used in the analysis

For other material analysis we used same procedure for new material. For that we select material from table and we do same FEA analysis on valve. getting require result.

Table. 2 Material composition & mechanical properties table:

ASTM DESIGNATION	CARBON STEEL		ALLOY STEEL				316 STAINLESS	
	A105	Class 1 A350 LF2	A182 F5	A182 F9	Class 2 A182 F11	Class 3 A182 F22	A182 F316	A182 F316L
Carbon, C	0.22 *	0.22 *	0.15	0.15	0.1-0.2	0.05-0.15	0.08	0.03
Silicon, Si	0.10-0.35	0.15-0.30	0.5	0.5-1.0	0.5-1.0	0.5	1.0	1.0
Manganese, Mn	0.60-1.05	0.60-1.35	0.3-0.6	0.3-0.6	0.3-0.8	0.3-0.6	2.0	2.0
Phosphorus, P	0.035	0.035	0.03	0.03	0.04	0.04	0.045	0.045
Sulfur, S	0.04	0.04	0.03	0.03	0.04	0.04	0.03	0.03
Nickel, Ni	0.4	0.4	0.5				10.0-14.0	10.0-15.0
Chromium, Cr	0.3	0.3	4.0-6.0	8.0-10.0	1.0-1.5	2.0-2.5	16.0-18.0	16.0-18.0
Copper, Cu	0.4	0.4						
Vanadium, V	0.08	0.08						
Columbium, Cb		0.02						
Molybdenum, Mo	0.12	0.12	0.44-1.00	0.9-1.1	0.44-1.00	0.87-1.13	2.0-3.0	2.0-3.0
Tensile, psi, min.	70,000	70,000-95,000	70,000	85,000	70,000	75,000	75,000	70,000
Yield, psi, min.	36,000	36,000	40,000	55,000	40,000	45,000	30,000	25,000
Elongation, %, min.	22	22	20	20	20	20	30	30
R. of Area %, min.	30	30-80	35	40	30	30	50	50
Hardness, HB	187	197	143-217	179-217	143-207	156-207		

V. CONCLUSION

After this experiments we can conclude by using this FEA analysis we get optimize material for ball valve and we can reduce the wall thickness for maximum condition, we also reduce the weight of valve and increase the strength of valve by using proper selection of materials. So ball valve life is increase and ball valve can sustain maximum condition of pressure and temperature.

REFERENCES

- [1] "Establishment of Relationships between Material Design and Product Design Domains by Hybrid FEM-ANN Technique" by K. Soorya Prakash, S. S. Mohamed Nazirudeen and M. Joseph Malvin Raj Department of Mechanical Engineering, Anna University Coimbatore, Tamil Nadu, India
- [2] "Development of validated pressure locking methodology for gate valve" By J.K.Wang, M.S.Kalsi, and S.S, Averitt Kalsi engineering, Inc.
- [3] "ANALYSIS AND DESIGN OF CRYOGENIC BALL VALVE" by Dong-Soo KIM\*, Myoung-Sub KIM, Roll-to-Roll PEMS Team, Nano Mechanical Systems Research Division, Korea Institute of Machinery & Materials 171, Jang-Dong, Yuseong-Gu, Daejeon, Korea
- [4] "The Effect of Thermal Cycling on Seals in Ball Valves and Clamp-Type Fittings" by Dave Simko, the Official Journal of ISPE.
- [5] "Design factors for "linear" ball valve: theoretical and experimental studies" by Thananchai Leephakpreeda Ph.D.(Mechanical Engineering), Assoc. Prof., School of Manufacturing Systems and Mechanical Engineering, Sirindhorn International Institute of technology, Thammasat University, Rangsit Campus, P.O. Box 22 Thammasat-Rangsit Post Office, Pathum Thani 12121, 14 September 2004
- [6] "EXPERIMENTAL RESIDUAL STRESS ANALYSIS OF WELDED BALL VALVE" By Pavel Macura 1, František Fojtík 2, Radomír Hrnčář Faculty of Mechanical Engineering, VŠB – TU Ostrava, Czech Republic.
- [7] " CFD Modeling of Globe Valves for Oxygen Application" by Aditi Oza1, Sudipto Ghosh2 and Kanchan Chowdhury1 1Cryogenic Engineering Centre, Indian Institute of Technology, Kharagpur 721 302, India, Department of Metallurgical & Materials Engineering, Indian Institute of Technology, Kharagpur 721 302, India
- [8] "Avoiding Pressure Surge Damage in Pipeline Systems" by Prof ARD Thorley, Prof Nakayama & RF Boucher.
- [9] "Using Computational Fluid Dynamics to Predict the Onset of Cavitation (xFz)" by Alan H. Glenn, Gifford Z. Decker, FLOWSERVE FCD Valtek Control Products.
- [10] "NUMERICAL ANALYSIS OF BUTTERFLY VALVE-PREDICTION OF FLOW COEFFICIENT AND HYDRODYNAMIC TORQUE COEFFICIENT" by Xue guan Song1, Young Chul Park Graduate student, songxguan@yahoo.com.cn Professor, parkyc67@dau.ac.kr CAE Lab, Department of Mechanical Engineering, Dong-A University, 840 Hadan-dong, Saha-gu, Busan 604-714.
- [11] "Finite Element Modeling and Optimization of a Robot Boomer" by Ulf Sellgren ,Department of Machine Design, Royal Institute of Technology, Stockholm, Sweden
- [12] "ANALYSIS THE EFFECT OF DIFFERENT TYPE VALVE USE AT DIFFERENT PRESSURE IN PIPING SYSTEM" by MUZAKKIR BIN SHUKRI, Faculty of Mechanical Engineering UNIVERSITI MALAYSIA PAHANG.

# Effect of solvents and their concentrations on extraction of Caffeine from different tea samples

A. Neha Patni, B. Nagja Tripathi, C. Viraj Baxi, D. Sakshi Chauhan.

**Abstract**—Caffeine occurs naturally in over 60 plant species including tea, coffee and cola. It is produced commercially both by extraction from natural sources and synthetic procedures. The majority of caffeine produced is used in the beverage industry. It is also used as a flavour enhancer in foods and as a flavouring agent in baked goods, frozen dairy desserts, gelatins, puddings, fillings and softcandy. Caffeine is also used therapeutically.

Data collected from a number of epidemiological studies on the long-term health effects of caffeine have indicated no marked elevations in risk for most diseases, including cardiovascular disease, ulcers, breast disease and for effects of various target organs. Absorption of caffeine from the gastrointestinal tract is pH-dependant, rapid and virtually complete. After absorption, it is distributed rapidly into the body fluids. It binds to the plasma proteins, mainly albumins. Metabolism is performed by hepatic microsomal enzymes and does not appear to occur significantly in other organs. In humans, caffeine is metabolized into more than 25 metabolites, primarily paraxanthine, theobromine and theophylline. In humans, acute exposure to caffeine can cause gastric symptoms, insomnia, diuresis, restlessness, headache and tremors. At concentrations of up to 10 mol/mL in blood, caffeine stimulates the central nervous system (CNS).

These days caffeine is being utilized as flavour-enhancer by most of the beverage (soft drinks), coffee and tea companies. So, its analytical estimation/determination methods become more important for its regulation. Herein, we have validated the extraction and TLC method, for the quick determination of caffeine from marketed samples of tea and coffee. The method not only finds importance in the quality control analysis of caffeine, but also shows the importance of screening marketed tea and coffee samples for the caffeine content.

**Index Terms**—Caffeine, Solvents, Tea extract, Chloroform

## I. INTRODUCTION

Caffeine is a bitter, white crystalline xanthine alkaloid that acts as a stimulant drug. Caffeine is found in varying quantities in the seeds, leaves, and fruit of some plants, where it acts as a natural pesticide that paralyzes and kills certain insects feeding on the plants. It is most commonly consumed by humans in infusions extracted from

the bean of the coffee plant and the leaves of the tea bush, as well as from various foods and drinks containing products derived from the kola nut. Other sources include yerba maté, guarana berries, guayusa, and the yaupon holly.

Caffeine was first isolated from coffee in 1820 by the German chemist Friedlieb Ferdinand Runge,<sup>[6]</sup> and then independently in 1821 by French chemists Pierre Robiquet, Pierre Pelletier, and Joseph Caventou. Pelletier coined the word "cafeine" from the French word for coffee (*café*), and this term became the English word "caffeine".

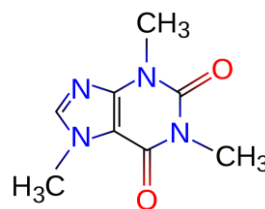


Fig. 1: Molecular structure of caffeine

**Significance of Caffeine Extraction:** These days caffeine is being utilized as flavour-enhancer by most of the beverage (soft drinks), coffee and tea companies. So, its analytical estimation/determination methods become more important for its regulation. Herein, we have validated the extraction and TLC method, for the quick determination of caffeine from marketed samples of tea and coffee. The method not only finds importance in the quality control analysis of caffeine, but also shows the importance of screening marketed tea and coffee samples for the caffeine content. Moreover, decaffeination serves the dual purpose of providing caffeine-free beverages on the one hand and caffeine, on the other, that finds several applications in fields such as medicine, sports etc.

## II. PROPERTIES OF CAFFEINE

Molecular formula	C <sub>8</sub> H <sub>10</sub> N <sub>4</sub> O <sub>2</sub>
Molar mass	194.19 g mol <sup>-1</sup>
Exact mass	194.080376 u
Appearance	Odorless, white needles or powder
Density <sup>[11]</sup>	1.23 g/cm <sup>3</sup> , solid
Melting point	227–228 °C, 500-501 K (anhydrous)

A:- Assistant Professor in Chemical Engineering Department, Nirma Institute of Technology, Nirma University, Ahmedabad, Gujarat, India. (Email id:- neha.patni @nirmauni.ac.in)

	234–235 °C, 507-508 K (monohydrate)
Boiling point	178 °C, 451 K, 352 °F (subl.)
Solubility in water	2.17 g/100 mL (25 °C) 18.0 g/100 mL (80 °C) 67.0 g/100 mL (100 °C)
Acidity (pK <sub>a</sub> )	–0.13–1.22
Dipole moment	3.64 (calculated)

### III. SOURCES AND CONSUMPTION

- Caffeine is found in many plant species, where it acts as a natural pesticide, with high caffeine levels being observed in seedlings till developing foliage but lacking mechanical protection; caffeine paralyzes and kills certain insects feeding upon the plant. High caffeine levels have also been found in the surrounding soil of coffee bean seedlings.<sup>[3]</sup>
- Common sources of caffeine are coffee, tea, and (to a lesser extent) chocolate derived from cocoa beans. Less commonly used sources of caffeine include the yerba maté, guarana and ilex guayusa plants, which are sometimes used in the preparation of teas and energy drinks. Two of caffeine's alternative names, *mateine* and *guaranine*, are derived from the names of these plants. One of the world's primary sources of caffeine is the coffee "bean" (which is the seed of the coffee plant), from which coffee is brewed. Caffeine content in coffee varies widely depending on the type of coffee bean and the method of preparation used; even beans within a given bush can show variations in concentration. In general, one serving of coffee ranges from 80–100 milligrams, for a single shot (30 milliliters) of arabica-variety espresso, to approximately 100–125 milligrams for a cup (120 milliliters) of drip coffee. In general, dark-roast coffee has slightly less caffeine than lighter roasts because the roasting process reduces a small amount of the bean's caffeine content. *Arabica* coffee normally contains significantly (+/-50%) less caffeine than the *robusta* variety. Coffee also contains trace amounts of theophylline, but no theobromine.
- Tea is another common source of caffeine. Although tea contains more caffeine than coffee (by dry weight), a typical serving contains much less, as tea is normally brewed much weaker. Besides strength of the brew, growing conditions, processing techniques and other variables also affect caffeine content. Certain types of tea may contain somewhat more caffeine than other teas. Tea contains small amounts of theobromine and slightly higher levels of theophylline than coffee. Preparation and many other factors have a significant impact on tea, and color is

a very poor indicator of caffeine content. Teas like the pale Japanese green tea, *gyokuro*, for example, contain far more caffeine than much darker teas like *lapsang souchong*, which has very little.

- Caffeine is also a common ingredient of soft drinks, such as cola, originally prepared from kola nuts. Soft drinks typically contain about 10 to 50 milligrams of caffeine per serving. By contrast, energy drinks, such as Red Bull, can start at 80 milligrams of caffeine per serving. The caffeine in these drinks either originates from the ingredients used or is an additive derived from the product of decaffeination or from chemical synthesis. Guarana, a prime ingredient of energy drinks, contains large amounts of caffeine with small amounts of theobromine and theophylline in a naturally occurring slow-release excipient.
- Chocolate derived from cocoa beans contains a small amount of caffeine. The weak stimulant effect of chocolate may be due to a combination of theobromine and theophylline, as well as caffeine. A typical 28-gram serving of a milk chocolate bar has about as much caffeine as a cup of decaffeinated coffee, although some dark chocolate currently in production contains as much as 160 mg per 100g.

### IV. USES OF CAFFEINE

- Caffeine is most commonly used to improve mental alertness, but it has many other uses. Caffeine is used by mouth or rectally in combination with painkillers (such as aspirin and acetaminophen) and a chemical called ergotamine for treating migraine headaches. It is also used with painkillers for simple headaches and preventing and treating headaches after epidural anaesthesia.
- Caffeine is also used for weight loss and type 2 diabetes.<sup>[21]</sup>
- Caffeine is one of the most commonly used stimulants among athletes. Taking caffeine, within limits, is allowed by the National Collegiate Athletic Association (NCAA). Urine concentrations over 15 mcg/mL are prohibited. It takes most people about 8 cups of coffee providing 100 mg/cup to reach this urine concentration.<sup>[1]</sup>
- Caffeine creams are applied to the skin to reduce redness and itching in dermatitis.
- Healthcare providers sometimes give caffeine intravenously (by IV) for headache after epidural anesthesia, breathing problems in newborns, and to increase urine flow.
- More controversial therapeutic uses of late the production of seizures during caffeine are these: to kill skin fungi; to enhance the toxic effects of chemicals used in cancer therapy; and to facilitate electroconvulsive therapy.



## V. DECAFFEINATION

Extraction of caffeine from coffee, to produce decaffeinated coffee, is an important industrial process and can be performed using a number of different solvents. Benzene, chloroform, trichloroethylene, and dichloromethane have all been used over the years but for reasons of safety, environmental impact, cost, and flavor, they have been superseded by the following main methods:

- **Water extraction**
  - Coffee beans are soaked in water.
  - The water, which contains many other compounds in addition to caffeine and contributes to the flavor of coffee, is then passed through activated charcoal, which removes the caffeine.
  - The water can then be put back with the beans and evaporated dry, leaving decaffeinated coffee with its original flavor.
  - Coffee manufacturers recover the caffeine and resell it for use in soft drinks and over-the-counter caffeine tablets.
- **Roselius process**
  - The first commercially successful decaffeination process was invented by Ludwig Roselius and Karl Wimmer in 1903. It involved steaming coffee beans with a brine (salt water) solution and then using benzene as a solvent to remove the caffeine..
  - Due to health concerns regarding benzene, this process is no longer used commercially.
- **Direct method**
  - In the direct method, the coffee beans are first steamed for 30 minutes and then repeatedly rinsed with either dichloromethane or ethyl acetate for about 10 hours.
  - The solvent is then drained away and the beans steamed for an additional 10 hours to remove residual solvent.
  - Sometimes coffees that are decaffeinated using ethyl acetate are referred to as naturally processed because ethyl acetate can be derived from various fruits or vegetables; but, because of the impracticality of gathering natural ethyl acetate, the chemical used for decaffeination is synthetic.
- **Indirect method**
  - In the indirect method, beans are first soaked in hot water for several hours, in essence, making a strong pot of coffee.
  - Then the beans are removed and either dichloromethane or ethyl acetate is used to extract the caffeine from the water.
  - As in other methods, the caffeine can then be

separated from the organic solvent by simple evaporation. The same water is recycled through this two-step process with new batches of beans.

- An equilibrium is reached after several cycles, where the water and the beans have a similar composition except for the caffeine.
- After this point, the caffeine is the only material removed from the beans, so no coffee strength or other flavorings are lost. Because water is used in the initial phase of this process, sometimes indirect method decaffeination is referred to as "water-processed" even though chemicals are used.
- **CO<sub>2</sub> process<sup>[5]</sup>**
  - This process is technically known as supercritical fluid extraction. In the carbon dioxide method, the caffeine is stripped directly from the beans by a highly compressed semi-liquid form of carbon dioxide.
  - Pre-steamed beans are soaked in a bath of supercritical carbon dioxide at a pressure of 73 to 300 atmospheres.
  - After a thorough soaking for around ten hours, the pressure is reduced, allowing the CO<sub>2</sub> to evaporate, or the pressurized CO<sub>2</sub> is run through either water or charcoal filters to remove the caffeine.
  - The carbon dioxide is then used on another batch of beans. This liquid works better than water because it is kept in supercritical state near the transition from liquid to gas, combining favorable diffusivity properties of the gas with increased density of a liquid.
  - This process has the advantage that it avoids the use of potentially harmful substances.
- **Triglyceride process**
  - Green coffee beans are soaked in a hot water/coffee solution to draw the caffeine to the surface of the beans.
  - Next, the beans are transferred to another container and immersed in coffee oils that were obtained from spent coffee grounds.
  - After several hours of high temperatures, the triglycerides in the oils remove the caffeine—but not the flavor elements—from the beans.

## VI. CHARACTERISATION OF CAFFEINE

The following tests can be used for the characterization of the caffeine extract:

- **Murexide Test**  
Caffeine when heated with HCl & potassium chlorate gives a residue which on exposure to ammonia vapours gives purple colour.

- **Thin Layer Chromatography**

Stationary phase: Thin layer of silica gel on a TLC plate.

Mobile phase: a. Ethyl acetate: Formic acid: Glacial acetic acid : water (50: 5.5: 5.5: 1)  
or b. Chloroform : ethanol (9.5: 0.5)

Mode of Detection: Spraying of modified Dragendorff's solution on the plate.

The presence of caffeine can be confirmed by comparing the value of the Retention Factor  $R_f$  of the sample with that of standard caffeine.

- **Tannic acid Test**

Solution of caffeine in water, gives a white precipitate on addition of Tannic acid.

- **Melting Point Characterisation of Caffeine salicylate**

Caffeine Salicylate is prepared by the acid-base reaction of Caffeine and Salicylic acid. The compound melts at  $137^{\circ}\text{C}$  which is the theoretical melting point of Caffeine Salicylate, thus confirming that the obtained compound is Caffeine.

## VII. EXPERIMENTAL PROCEDURE<sup>[22]</sup>

### PURPOSE

- To carry out an extraction of an aqueous solution of coffee with an organic solution.
- To isolate caffeine from coffee.
- Characterization of Caffeine

### MATERIALS NEEDED

- 10 g Tea leaves powder
- 100 mL distilled water
- 5 g solid sodium carbonate
- 75 mL solvent (Chloroform/Ethyl Acetate/2-Propanol)
- 250 mL and 500mL beaker
- glass funnel
- 50 mL beaker
- Glass separating funnel

### PROCEDURE

Extraction of the Tea Solution

1. Prepare Tea extract by boiling tea leaves in distilled water for 30 mins and filter



Fig. 6: Tea sample

2. Clean 500 mL flask, add 100 mL of tea extract.



Fig. 7: Tea extract

3. Add approximately five grams of sodium carbonate ( $\text{Na}_2\text{CO}_3$ ) to the tea solution and stir well and filter. This will remove the tannins present in tea extract.
4. Take the filtrate in a glass separating funnel and extract with three successive quantities of 25 mL of organic solvent like chloroform by shaking well. Do not shake the mixture vigorously or an emulsion will form.
5. Allow the mixture to stand and separate into two layers; a dark aqueous top layer and a clear organic bottom layer.



Fig. 8: Mixture in separator

6. Separate the organic layer by opening the cork.
7. Combine the organic layer and evaporate to dryness to obtain a crude residue of caffeine

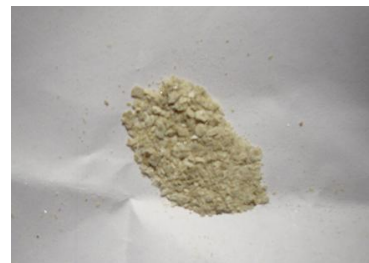


Fig. 9: Pure caffeine

## VIII. RESULTS

- Solvent Testing:  
Sample : 100 ml tea extract

SOLVENT	CAFFEINE EXTRACTED (g)	YIELD (%)
Chloroform	1.2	96
2-Propanol	0.9	72
Ethyl Acetate	0.75	60

Yield percentage =  $\frac{\text{Amount of Caffeine Extracted (grams)}}{\text{Amount of caffeine present in the sample (grams)}} * 100$   
 Caffeine initially present in the sample = 0.125g

- Sample : 100 ml tea extract  
 Solvent : 25 ml Chloroform

The extraction procedure was carried out in three stages. The tea extract obtained as the upper layer after taking out chloroform from the separating funnel (1<sup>st</sup> stage); is again mixed with 25ml of chloroform.

Sample	1 <sup>st</sup> Stage Extraction (g)	2 <sup>nd</sup> Stage Extraction (g)	3 <sup>rd</sup> Stage Extraction (g)	Total (g)
Tata Agnee	0.06	0.03	0.01	0.1
Red Label	0.08	0.07	Negligible	0.15
Waghbakri	0.07	0.04	0.01	0.12
Mili CTC	0.06	0.05	Negligible	0.11
Lipton	0.05	0.03	0.01	0.09

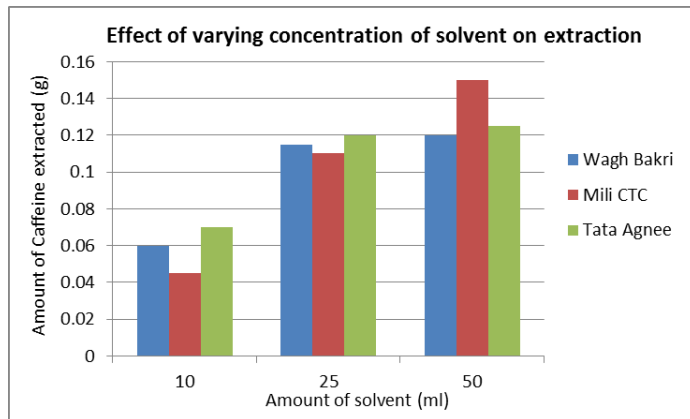
- Effect of varying concentration of solvent on extraction

The amount of chloroform used in the extraction procedure was varied to observe the effect on yield. This was done using 10ml, 25ml and 50 ml of chloroform.

It was observed that the yield of caffeine increased with increase in solvent concentration.

Sample	9.09% chloroform	50% chloroform	33.33% chloroform
Wagh Bakri	0.060g	0.115g	0.12g
Mili CTC	0.045g	0.11g	0.15g
Tata Agnee	0.07g	0.12g	0.125g

Sample = 100 ml tea extract



IX. CONCLUSIONS

Chloroform gives the maximum yield of caffeine at given conditions of temperature, pressure and concentration. This is due to its low miscibility in water compared to other solvents. However, 2-Propanol is used industrially due to the harmful nature of chloroform to environment. The effect of varying the concentration of the solvent shows that as concentration increases, yield also increases.

X. ACKNOWLEDGEMENT

The authors acknowledge the Nirma University for providing the financial support for the conduction of experiment.

XI. REFERENCES

- [1] Armstrong LE, Casa DJ, Maresh CM, Ganio MS (July 2007). "Caffeine, fluid-electrolyte balance, temperature regulation, and exercise-heat tolerance". *Exerc Sport Sci Rev* 35 (3): 135–40.
- [2] Anvar Shalmashi, Mohammad Abedi, Fereshteh Golmohammad, Mohammad Hassan Eikani, "Isolation Of Caffeine From Tea Waste Using Subcritical Water Extraction", *Journal of Food Process Engineering*, Volume 33, Issue 4, pages 701–711, August 2010.
- [3] Baumann, T. W. (1984). "Metabolism and excretion of caffeine during germination of *Coffea arabica* L". *Plant and Cell Physiology* 25 (8): 1431–6.
- [4] Bolton, Sanford (1981). "Caffeine: Psychological Effects, Use and Abuse", *Orthomolecular Psychiatry* 10 (3): 202–211.
- [5] C.B. Mehr, R.N. Biswal, J.L. Collins, H.D. Cochran, "Supercritical carbon dioxide extraction of caffeine from guarana", *The Journal of Supercritical Fluids*, Volume 9, Issue 3, September 1996, Pages 185–191
- [6] Ferdinand Friedlieb, Runge (1820). *Neueste phytochemische Entdeckungen zur Begründung einer wissenschaftlichen Phytochemie [Latest phytochemical discoveries for the founding of a scientific phytochemistry]*. Berlin: G. Reimer. pp. 144–159.
- [7] Fisone G, Borgkvist A, Usiello A. "Caffeine as a psychomotor stimulant: mechanism of action". *Cell. Mol. Life Sci.* 61 (7–8): 857–72, April 2004
- [8] H Misra, D Mehta, BK Mehta, M Soni, DC Jain, "Study of extraction and HPLC - UV method for estimation of caffeine in marketed tea (*Camellia sinensis*) granules", *Year 2009, Volume 3, Issue 1, Page 47-51*
- [9] Hammami, M. M., Al-Gaai, E. A., Alvi, S., & Hammami, M. B. (2010). "[Interaction between drug and placebo effects: A cross-over balanced placebo design trial]". *Trials* 11 (110): 1–10.
- [10] Hampp, A., *J. Chem. Educ.*, 1996 (73) 1172.

- [11] International Occupational Safety and Health Information Centre (CIS), "Caffeine".
- [12] Lovett, Richard (24 September 2005). "Coffee: The demon drink?". *New Scientist* (2518). Retrieved 2009-08-03.
- [13] M Paridhavi, S.S. Agrawal, "Herbal Drug Technology", University's Press India pvt lmt, 2007.
- [14] McArdle, William (2010). *Exercise Physiology*. 7th edition. Baltimore, MD: Lippincott Williams and Wilkins. p. 559
- [15] Medscape Multi-Drug Interaction Checker, "Drug Interaction: Caffeine Oral and Fluvoxamine Oral".
- [16] Murray, D.S.; Hansen, P.J., *J. Chem. Educ.*, 1995 (72) 851
- [17] Muthanna J Mohammed, Firas A Al-Bayati, "Isolation, identification and purification of caffeine from *Coffea Arabica* L. and *Camellia sinensis* L.": A combination antibacterial study January 2009, Volume , Issue, 52-57
- [18] Nehlig, A; Daval, JL; Debry, G (1992). "Caffeine and the central nervous system: mechanisms of action, biochemical, metabolic and psychostimulant effects". *Brain Research Reviews* 17 (2): 139–70.
- [19] Ribeiro JA, Sebastião AM (2010). "Caffeine and adenosine". *J. Alzheimers Dis.* 20 Suppl 1: S3–15.
- [20] Tadelech Atomssa and A.V. Gholap, Characterization of caffeine and determination of caffeine in tea leaves using uv-visible spectrometer, *African Journal of Pure and Applied Chemistry* Vol. 5(1), pp. 1-8, January 2011.
- [21] Van Dam RM (2008). "Coffee consumption and risk of type 2 diabetes, cardiovascular diseases, and cancer". *Applied physiology, nutrition, and metabolism* 33 (6): 1269–1283.
- [22] <http://www.seriaz.org/downloads/4-caffiene.pdf>

# A Study on MAP process at laboratory scale for the removal of $\text{NH}_4\text{-N}$

A K J Suthar, B N P Chokshi

**Abstract--**This laboratory scale study was conducted on the MAP process to improve struvite precipitation and more efficient removal of  $\text{NH}_4\text{-N}$  from the wastewater. MAP process is one of the innovative Physico-chemical processes which can majorly contribute for the removal of  $\text{NH}_4\text{-N}$  by forming the struvite complex under required condition. In order to meet stringent discharge standards for  $\text{NH}_4\text{-N}$ , an experimental work was carried out focusing the general problem of effluent discharge for the industries of India containing high  $\text{NH}_4\text{-N}$  stream. For this we evaluated the effort of feeding sequence of precipitating reagent (lime, flocculent, magnesium and phosphate) for  $\text{NH}_4\text{-N}$  removal by forming MAP i.e. struvite deposit keeping calcium interference in mind. Various studies reviewed which effectively precipitate struvite by an addition of excess magnesium and phosphate sources with different stoichiometric ratios. The stoichiometric calculation for the optimized dosage of reagent makes the precipitation more efficient. Different stoichiometric ratio were tested to determine the required  $\text{Mg}^{2+}$  and  $\text{PO}_4^{3-}$  concentration for maximum  $\text{NH}_4\text{-N}$  removal by keeping  $\text{NH}_4^+$  constant for the effluent and kinetics of struvite formation was studied.

**Index Terms--** Ammoniacal Nitrogen, struvite precipitation, MAP

## I. INTRODUCTION

Ammoniacal nitrogen is a measure for the amount of ammonia a toxic pollutant often found in landfill leachate and in waste products, such as sewage, liquid manure and other liquid organic waste products. Large scale urbanization, a consequence of economic development is leading to production of huge quantities of waste water in India and posing serious environmental problems for their disposal. It is evident that the protection of our water resources is of major importance on a global scale. The treatment and disposal of sludge produced during waste treatment is one of the most critical environmental issues of today. Sludge produced is large in volume and hazardous. Another issue of concern is that the sewage sludge produced and effluents are frequently disposed off on agricultural lands as fertilizer and irrigation purpose, respectively, due to their nutrient contents, especially N and P without any treatments, but they may induce plant and soil toxicity and may have depressive effects on the metabolism of soil microorganisms[1,2]. Therefore, there is a need for ecologically sound technologies which are not only cost-effective, but also sustainable in terms of possible recovery of recyclable constituents from sewage sludge as they are rich in nutrients and have higher organic content.  $\text{NH}_4\text{-N}$  has been identified as one of the major toxicants to

microorganisms in the treatment system, suggesting that pre treatment prior to the biological treatment system is required to reduce the concentration of  $\text{NH}_4\text{-N}$  [3]. Ammoniacal

Nitrogen contamination is highly toxic and it would adversely affect aquatic life if is discharged above sustainable limit.  $\text{NH}_4\text{-N}$  is the toxicant that is causes the death of tilapia fish [4]. The high content of ammoniacal nitrogen is the major factor that affects the toxicity of wastewater.

### 1) Sources of $\text{NH}_4\text{-N}$

The removal of dyes from industrial effluents is an area of research receiving increasing attention as government legislation surrounding the release of contaminated effluent is becoming increasingly stringent. The presence of very low concentrations of dyes in effluent is highly visible and undesirable. The effluents from the dyes industry containing high concentration of  $\text{NH}_4\text{-N}$  contribute majorly to the wastewater pollutants. There are more than 100,000 dyes available commercially, most of which are difficult to decolorize due to their complex structure and synthetic origin. These systems depended on biological activity and were mostly found inefficient in the removal of the more resistant synthetic dyes.

Most wastewater treatment plants do not accept dye-containing effluent due to its adverse effects on microbial populations that would affect the biological treatment i.e. aeration tank, secondary clarifier. Majorly fertilizer and dyes manufacturing industries have high  $\text{NH}_4\text{-N}$  content in their effluent. The main pollutant in nitrogenous fertilizer manufacture is ammoniacal nitrogen [5]. Therefore the treatment of effluent for removal recovery/recycle for ammoniacal nitrogen is essential.

### 2) Alternates for the removal of $\text{NH}_4\text{-N}$

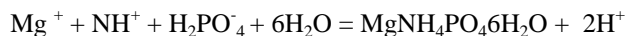
Many processes have been studied for exploring the treatment of wastewater containing ammoniacal nitrogen. Of these few processes are of commercial importance and may be opted for treatment. The available processes of practical importance can be broadly divided into two main categories: physico chemical processes which involve stripping of ammonia of effluent by air or stream and ion exchange [6]. It also includes the MAP process which satisfies technical-economic feasibility. Other category is the biological processes which include nitrification followed by denitrification of ammonia and algal uptake of ammoniacal

nitrogen [6]. Some other methods of treatment which has less practical importance are chlorination, electrodialysis, reverse osmosis, distillation etc. Stripping of ammonia in wastewater is transferred to surrounding atmosphere, where as ion exchange capital and operating cost is very high.

### 3) Technical perspective of MAP process

Nitrogenous compound may be present in the wastewater effluent as ammoniacal nitrogen, organic nitrogen i.e. urea and oxidized nitrogen i.e. nitrate and nitrite. The total nitrogen consists of the inorganic forms of nitrogen (ammonium (NH<sub>4</sub>-N) and nitrate (NO<sub>3</sub><sup>-</sup>-N) and organic nitrogen (N<sub>org</sub>). The total kjeldahl nitrogen value provides combined values of ammoniacal nitrogen and organic nitrogen but not oxidized nitrogen. Most of the toxic effluent from dyes and fertilizer industries having high NH<sub>4</sub>-N is acidic and this in general are neutralized using lime slurry which brings calcium interference in the picture. As an alternative to eliminate high level of NH<sub>4</sub>-N in leachate, the precipitation of NH<sub>4</sub>-N by forming magnesium ammonium phosphate (Struvite) has been studied, demonstrating that struvite precipitation is an excellent pretreatment process [1]. Struvite crystallizes as a white orthorhombic crystalline structure, which is composed of magnesium, ammonium, and phosphate in equal molar concentrations [7]. Struvite precipitation is controlled by pH, supersaturation, temperature, and impurities such as calcium. Struvite solubility decreases with increasing pH, while above a pH of 9 its solubility begins to increase [8]. Struvite precipitation will not occur without a nucleation in itself. Struvite precipitation is principally based on the thermodynamic equilibrium of constituent ions in the solution, this study was conducted in the assumption that a performance of struvite precipitation might vary according to feeding sequence of chemicals of magnesium and phosphate source [9]. Considering the case of Indian industries Phosphate is limiting and hence being added last. MAP complex process is pH sensitive method. A complex becomes insoluble at pH above 8.5 and hence accurate control of pH has to be shown in operation. The modelling is complex because of the interference of various different possible magnesium and calcium mineral precipitations (ion competition and simultaneous precipitation). Dosing of Mg and P source

depends on the stoichiometric calculation. The molar ratio of Mg:N:P varies for different samples. The struvite process removes ammonia in a dewatered filtrate of a sludge treatment facility by struvite preparation reaction, and at the same time recovers the struvite prepared. According to the presented ionic reaction stoichiometric



The kinetics of homogeneous chemical reactions can be written with respect to one species as shown in equation (1).

$$-d[\text{C}]/dt = k[\text{C}]^n \quad (1)$$

Where C is the concentration of reactant, k is the rate constant and n is the order of reaction. If the equation is integrated for the first, second, and third order, it yields the following integrated equations respectively as shown is equation (2),(3) and (4).

$$\ln[\text{C}] = \ln[\text{C}_0] - kt \text{ for the first order} \quad (2)$$

$$1/[\text{C}] = 1/[\text{C}_0] + kt \text{ for the second order} \quad (3)$$

$$1/2[\text{C}]^2 = 1/2[\text{C}_0]^2 + kt \text{ for third order} \quad (4)$$

The formation of struvite crystals significantly reduced gaseous loss of ammonia and resulted in substantial increase in the ammonia content in the compost, attaining 1.5% [10]. In this context, the present study was conducted to determine the optimal doses of Mg and P salts for struvite crystallization keeping calcium interference in mind. Various molar ratios are defined in different literature i.e. maximum of 88% ammonia removal at a pH of 9.5 with added magnesium and phosphate to achieve an ammonium: magnesium: phosphate molar ratio of 1:1.25:1[11]. Hence it becomes essential to optimize the dosage. Depending on the composition of the wastewater, struvite precipitation can be used to remove ammonia (NH<sub>4</sub><sup>+</sup>), phosphate (PO<sub>4</sub><sup>3-</sup>) or both. Regardless of the compound that is targeted for removal, all of the studies up to date have utilized the addition of Mg<sup>2+</sup> ion, the usual limiting reactant in the formation of struvite, as the means for altering the solubility product equilibrium and initiating precipitation. Table-1 Shows the data on the removal of ammoniacal nitrogen by struvite precipitation process for different wastewater along with optimized pH and molar ratios used for different studies.

TABLE 1.  
DATA ON THE REMOVAL OF NH<sub>4</sub><sup>+</sup> AND PO<sub>4</sub> BY STRUVITE PRECIPITATION FOR DIFFERENT STOICHIOMETRIC RATIOS

Type of effluent.	Removal	Molar ratios	pH	% removal	Ref
municipal landfill leachates	NH <sub>4</sub> -N	(Mg:N:P) 1:1:1	9.0	85	[12]
Piggery wastewater	PO <sub>4</sub> -P	(Mg:N) 2.5:1	8.5	96	[13]
municipal landfill leachate	NH <sub>4</sub> -N	(Mg:N:P) 1:1.2:1.2	9.0	90	[8]
Anaerobically Treated wastes	NH <sub>4</sub> -N	(Mg:N:P) 1:1.25:1	9.5	88	[11]
semiconductor wastewater	NH <sub>4</sub> -N	(Mg:N:P) 1.2:1:1	9.2	89	[14]
anaerobic digester effluents	NH <sub>4</sub> -N	(Mg:N:P) 1.2:1:1.2	8.5	77	[15]
Combined wastewater from bovine and leather tanning factories	NH <sub>4</sub> -N	(Mg:N:P) 1:1:1	9.0	82	[16]
Effluent from the biologically treated opium alkaloid wastewater	NH <sub>4</sub> -N	(Mg:N:P) 1:1:1	9.2	65	[17]
Effluent from anaerobic treatment of wastewater + 2% leachate	NH <sub>4</sub> -N	(Mg:N:P) 1:1:1	9.2	77	[18]

## II. EXPERIMENTAL STUDY

It is carried to remove metal compounds from waste water. It is a two step process. In the first step precipitants are mixed with wastewater allowing the formation of insoluble metal precipitants. Initially all the parameters are estimated that are listed in table-2.

TABLE 2  
INITIAL PARAMETERS

pH	10.25
TDS	1.75ppt
NH <sub>4</sub> -N	1737mg/l
PO <sub>4</sub> <sup>-3</sup>	2.20mg/l

### 1) Experiment

The experiment is carried out in a 2L of beaker with moderate speed agitator. If the sample is acidic, it has to be neutralized using 10% lime slurry. The sample considered here is highly alkaline hence direct treatment can be done. Once all the parameters are estimated, calculative amount of precipitating reagents can be added. Other than the chemical other important thing in chemical precipitation is pH. Metal hydroxides are amphoteric in nature and can react chemically as acids or bases and their solubility increases towards higher or lower pH. Thus, there is an optimum pH for hydroxide precipitation for each metal.

Here pH should be maintained at about 8.5 for struvite precipitation. The sequential addition of reagent should be magnesium source followed by phosphate and then allowing the mixture to agitate for about and 30 minutes at moderate speed and estimated the results. Similarly, duplicated the procedure for reaction time of 60, 90 and 120 minutes of reaction time and estimated the result. Degree of agitation along with pH plays an important role in efficient formation of struvite. After two hour of continuous stirring, precipitant is allowed to settle down the container. For rapid settling, flocculent like poly electrolyte can be used. The supernatant is collected and again all the parameters are estimated to find out the parentage reduction in NH<sub>4</sub>-N.

### 2) Precipitating reagent

MgO: The use of magnesium oxide as a reagent for struvite precipitation is liable to result in the presence of free ammonia, because its poor solubility reads to molar overdosing and thus to pH increases. For the source of magnesium, other compounds like MgSO<sub>4</sub>.7H<sub>2</sub>O and others can also be added but magnesium oxide is cheaper than other options. Hence MgO for Mg source with phosphorus addition to cause struvite precipitation could be a viable solution for ammonium removal from sludge liquors. This would avoid the return of ammonium to the further stage, thus improving biological nutrient removal. Sodium Tripoly

Phosphate (STPP): It is easily soluble in water; it has got salient chelating capacity to ions of Ca and Mg etc. and can soften hard water to make suspension solution become into clear solution; it has got weak alkalinity but no corrosiveness. It is a surfactant, and has got outstanding emulsification to lubricants and fat.

### 3) Stoichiometric Calculation

It is necessary to optimize the dosage. The sample considered here has NH<sub>4</sub>-N of about 1800mg/l and the optimum ratio considered is 0.06:1:0.07. The calculation is shown in Table-3

TABLE-3  
CALCULATION OF PRECIPITATING REAGENT

NH <sub>4</sub> -N present in the sample	1.8/14 = 0.128gmole
Mg required to treat the sample,	0.128 x 0.06 = 7.71 x 10 <sup>-3</sup> gmole
Mg(24) content in MgO(40)	60%
70% MgO required,	(7.71 x 10 <sup>-3</sup> x 25)/0.6 x 0.7 = 0.4591g/l
STPP required	=(0.128 x 0.07x93) / 0.25 = 3.34g/l

### 4) Analytical sampling of NH<sub>4</sub>-N

The standard method for the detection of NH<sub>4</sub>-N is kjeldahl method though other method can be used like nessler's reagent method. 100ml of water along with 10ml of sample are added in round bottom flask and pH is maintained by addition of caustic solution. pH should be about 10.25 as absorption takes place at corresponding pH. Phenolphthalein as an indicator is added and the assembly is arranged for the simple distillation as shown in Figure-1.

Distillate is collected in a beaker containing absorbent i.e. boric buffer and indicator. 100ml of distillate is collected and then titrated against 0.1N H<sub>2</sub>SO<sub>4</sub>.

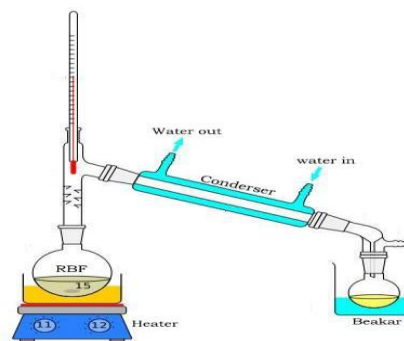


Fig.-1 Analytical Sampling of NH<sub>4</sub>-N

III. RESULT AND DISCUSSION

Models are important tools for process control. In the case of struvite, the modelling is complex because of the interference of various different possible magnesium and calcium mineral precipitations (ion competition and simultaneous precipitation). The recovered struvite contains hardly any toxic substance, and can be sold to fertilizer companies as a fertilizer raw material. Ammonium nitrogen is absorbed by soils but is largely converted to the nitrate form in a matter of one to three weeks in many soils. The result of various experimental runs is listed in Table-4. The

TABLE-4  
RESULT SHOWING PERCENTAGE REDUCTION

Initial NH <sub>4</sub> -N (mg/l)	NH <sub>4</sub> -N after treatment(mg/l)	% reduction
1768	772	56.33
1792	784	56.25
1686	752	55.39
1736	764	56.00
1698	756	55.47
Avg 1736	765	55.88

optimized ratio of Mg:N:P can be estimated by performing various experimental runs. The optimized ratio here is 0.06:1:0.07 as there is no prominent increase in the percentage reduction. For different influent sample, this ratio has to be optimized for better recovery. If high ratios without optimization are used directly for ammoniacal nitrogen recovery, it would be waste of chemicals. If influents are segregated for separate treatment, the recovery would be higher. Figure-2 shows the comparison between different stoichiometric ratios.

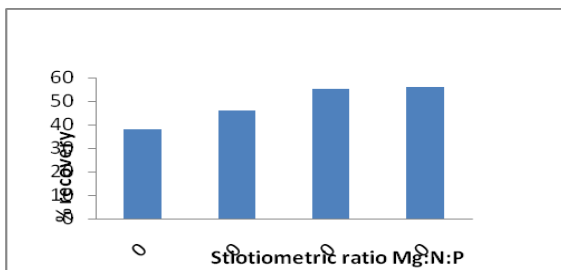


Fig. 2 Comparison of various ratios for 1hr. reaction time

The increase in the ratio would be the wastage of excess chemicals and would just add into the sludge formation, may even spoil the basic chemistry. If the sample is acidic, it becomes essential to neutralize the sample which would aid into interference of calcium. The comparison between

acidic and the basic samples after treatment is shown in Figure-3

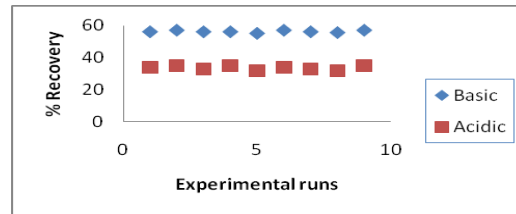


Fig-3 Comparison between acidic and basic samples after treatment for 1hr reaction time

The acidic sample shows less reduction because of calcium interference. For the molar ratio Ca:Mg above 1:1, calcium ions compete with magnesium to form an amorphous calcium phosphate changing the basic chemistry needed for MAP complex process and inhibiting struvite formation. In common effluent treatment if influent flow from equalization is arranged such that pH is adjusted about or above 7 without addition of lime, calcium interference would not be a major problem and reduction of NH<sub>4</sub>-N would be high. This is the general problem with acidic wastewater samples which have to be neutralized for the treatment. The time course of the reaction is given in Table-5. The reaction of struvite formation was considered to be homogeneous and crystal formation was not taken into account.

TABLE-5

KINETIC DATA (AVERAGE OF FIVE EXPERIMENTAL RUNS)

Time (min)	NH <sub>4</sub> -N (mg/l)	NH <sub>4</sub> -N (M)	ln[NH <sub>4</sub> -N]	1/[NH <sub>4</sub> -N]	1/2[NH <sub>4</sub> -N] <sup>2</sup>
0	1736	0.124	-2.0874	8.064	0.00768
30	1514	0.108	-2.2256	9.259	0.0058
60	1108	0.0791	-2.537	12.64	0.00312
90	822	0.058	-2.8473	17.24	0.00168
120	765	0.0546	-2.9077	18.31	0.0014

The time course of the reaction is given in Table-5 as the average of five sets of experimental data. The reaction was completed in about 120minutes. Integrated forms of the first, second, and third order reaction models were fitted to the experimental data. These are presented in Figure-3, 4 and 5 respectively.

The first order approach did not give satisfactory fit to the experimental data though R-square value of 0.96 was reasonable. Also third order did not fit well and gave poor R-square value of 0.93. The best fit was obtained with the second order rate approach with the R-square value of 0.96 which is reasonably good.



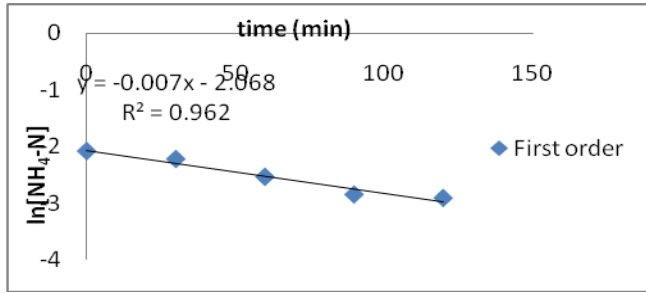


Fig. 4 Plot of ln[NH<sub>4</sub>-N] vs. time for first order reaction

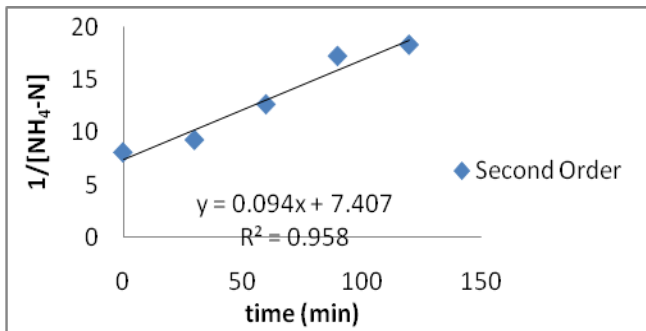


Fig. 5 Plot of  $1/[NH_4-N]$  vs.  $1/[NH_4-N]$  time for second order reaction

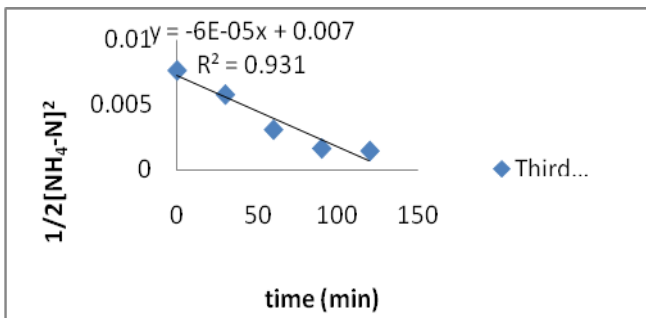


Fig. 5 Plot of  $1/2[NH_4-N]^2$  vs.  $1/[NH_4-N]$  time for second order reaction

The constants of the kinetic model thus obtained from the plots are presented in Table 6.

TABLE-6  
ESTIMATION OF KINETIC PARAMETERS IN KINETIC MODEL

Order	R <sup>2</sup>	K
1 <sup>st</sup>	0.962	0.4h <sup>-1</sup>
2 <sup>nd</sup>	0.958	5.64 l/mol h
3 <sup>rd</sup>	0.931	3.6x10 <sup>-3</sup> l <sup>2</sup> /mol <sup>2</sup> h

We have observed much faster precipitation kinetics completed in 120minutes. In the literature, there are a few reports dealing with struvite kinetics. Different in the result is observed and is tabulated. Different process with different pH

and order of reaction is tabulated. Different results were compared in Table 7.

TABLE-7  
KINETIC PARAMETER OBSERVED IN VARIOUS LITERATURE

order	k	pH	Ref.
2,3	-	8.0	[19]
1	4.2 h <sup>-1</sup>	-	[20]
1	3.7 h <sup>-1</sup>	8.4	[21]
1	7.9 h <sup>-1</sup>	8.7	[21]
1	12.3 h <sup>-1</sup>	9.0	[21]
2	5.64 l/mol h	8.5	Present study

The difference between our results and the others might be due to the fact that we have studied nucleation process whereas in literature [19]-[21], have investigated crystal growth, both processes are governed by different mechanisms resulting in different kinetic expressions.

#### IV. CONCLUSIONS

The process reduces NH<sub>4</sub><sup>+</sup>-N from high ammoniacal streams along with reduction in COD of around 18%. The precipitated struvite has shown to have relatively low heavy metal content, compatible with re-use as a fertilizer. Re-use routes are currently being developed as an organic horticultural fertilizer ingredient. The ongoing experiments showed the systematic pH controlled MAP process which is feeding an increasingly detailed data bank concerning struvite precipitation representative of waste water. Phosphate, magnesium, ammonium ions are maintained at optimized ratios for efficient precipitation. The ratio for different influent is different and has to be optimized for better recovery. And second order kinetic model was developed to predict the rate and extent of NH<sub>4</sub>-N removal. This method converts the hazardous waste into non-hazardous useful nutritious resource. MAP process can stand tall for green revolution.

#### V. ACKNOWLEDGEMENT

I am grateful to the department of chemical engineering, Nirma University which facilitated the completion of the presented experimental work.

#### VI. REFERENCES

- [1] Ayusho M, Pascual J.A, Garcia. A, Hremendez. T, "Evaluation of urbanwastes for agricultural use", *Soil Sci. Plant Nutr.*, vol. 42, 1996, pp 105-111
- [2] Ueno. Y, Fujii M., "3 years operates experience selling recovered struvite from full-scale plant", *Tech. Biotech.* 50, 1996, pp 257-264

- [3] Nigam. P, Armour G., Banat I.M, Singh D, Marchant R., “Physical removal of textile dyes from effluents and solid-state fermentation of dye-adsorbed agricultural residues”, *Bio. Tech.* 72, 2000, pp 219-226
- [4] “Advanced Biological treatment processes” By Lawrence K. Wang, Norman C. Pereira, Yung-Tse Hung, *Wiley inc. pub.*, edition 1, pp 345
- [5] “Waste minimization and cost reduction for the process industries” By Paul N. Cheremisinoff, (treatment of effluent fertilizer industry example), edition 1, pp 285-299
- [6] Lee S.I, Weon S.Y, Lee C.W, Koopman. B, “Removal of nitrogen and phosphate from wastewater by addition of bittern”, *Chemosphere* 51, 2003, pp 265–271
- [7] Michalowski T, Pietrzyk A, “A thermodynamic study of struvite plus water system”, *Jr. of Talanta* 68, 2006, pp 594–601
- [8] Daekun K, Ryu H.D, Kim M.S, Kim J., Lee S., “Enhancing struvite precipitation potential for ammonia nitrogen removal in municipal landfill leachate”, *Jr. of Env. Quality* 30, 1996, pp 1548-1563
- [9] Khwairakpam M, Bhargava .R, “Vermitechnology for sewage sludge recycling”, *Jr. of Hrd. Mat.* 161, 2009, pp 948–954
- [10] Jeong, YK, Hwang S.J, “Optimum doses of Mg and P salts for precipitating ammonia into struvite crystals in aerobic composting”. *Bio. Tech.* 96, 2004, pp 1–6
- [11] Miles A, Ellis T.G, “Struvite Precipitation potential for nutrient recovery from anaerobically treated wastes”. *Ames, IA*, 1996, pp. 50011-3232
- [12] Ozturk J, Altinbas M, Koyuncu I, Arikan O and Gomec I.Y, “Advanced physico-chemical treatment experiences on young municipal landfill leachates”, *waste management.* 23, 2003, pp 441-446
- [13] Huang H, Xu C and Zhang W, “Removal of nutrients from piggery wastewater using struvite precipitation and pyrogenation technology”, *Bio. Tech.*, Volume 102, Issue 3, 2011, pp.2523-2528
- [14] Ryu H.D, Kim D, Lee S, “Application of struvite precipitation in treating ammonium nitrogen from semiconductor wastewater”, *Jr. of Haz. Mat.* 156, 2008, pp.163–169
- [15] Turker M, Celen I, “Removal of ammonia as struvite from anaerobic digester effluents and recycling of magnesium and phosphate”, *Bio. Tech.* 98, 2007, pp. 1529–1534
- [16] Tunay O, Kabdasli I, Orhon D, Kolcak S. “Ammonia removal by magnesium ammonium phosphate precipitation in industrial wastewaters”, *Water Sci Tech.* 36, 1997, pp. 225–228
- [17] Altinbas M, Ozturk I, Aydin AF, ”Ammonia recovery from high strength agro-industry effluents” *Water Sci Tech.* 45, 2002, pp. 189–96
- [18] Altinbas M, Yangin C, Ozturk I. “Struvite precipitation from anaerobically treated municipal and landfill wastewaters” *Water Sci Tech.* 46, 2002, pp. 271–278
- [19] Gunn, D.J.,”Mechanism for the formation and growth of ionic precipitates from aqueous solutions in far”. *Chem. Soc.*, 1996, pp. 133– 140
- [20] Ohlinger, K.N., Young, T.M., Schroeder, E.D., “Post digestion struvite precipitation using a fluidized bed reactor”. *Jr. Environ. Eng.* 126, 2000, pp. 361–368
- [21] Nelson, N.O., Mikkelsen, R.L., Hesterberg, D.L., “Struvite precipitations in anaerobic swine lagoon liquid: effect of pH and Mg:P ratio and determination of rate constant”. *Bio. Tech.* 89, 2003, pp. 229–236

# Physicochemical and Biochemical Quality Assessment of Groundwater of Village Atul: Gujarat.

A. Bhadreshkumar. R. Sudani, B. Navrang B. Sudani, c. Hitarth. K. Soni

**Abstract—** The present work of this paper examines the effects of environmental pollution on the quality of groundwater of village Atul, district Valsad. Total 17 samples were collected from near by Atul area and were analyzed for water quality parameters like pH, Turbidity, Total Hardness, Total Alkalinity, Total Solid, Total Dissolved Solid, Nitrate, Phosphate, Fluoride, Chloride, Sulphate, BOD, COD, Total coliform, Faecal coliform . The result of analysis reveals that the groundwater of the area needs some degree of treatment before direct consumption as potable or house hold uses.

*Index Terms:* Atul, Groundwater, Quality, Pollution

## I. INTRODUCTION

Water is the life of all living organisms on this earth. Now a day due to industrial development most of the natural resources of water are being contaminated by one or other ways. Groundwater which uses for domestic and industrial water supply or irrigation is vital to whole mankind. According to WHO organization, about 80% of all the diseases in human beings are caused by water. In nature 97.2 % water is salty and 2.8 % water is fresh water, in which 20% water sources are from groundwater systems [1].

Village Atul is having a giant dyes and chemical manufacturing company Atul LTD. near by it. It is situated on the bank of river Par in district Valsad, state Gujarat, India.

The health of human being is closely related with the groundwater quality, if the quality of ground water becomes poor by excessive application of fertilizers or unsanitary conditions it must be very dangerous. Once the groundwater becomes polluted, it might be hard to stop the pollution and

becomes more imperative to monitor the quality of groundwater regularly [2].

## Study Area:

Atul is an industrial village like town place developed due to the company Atul Ltd, located in District Valsad, which is located at the Southernmost tip of Gujarat, The town is located on 72.940002° (10-mode) -72°56'24 " (60-mode) Longitude east and 20.547962° (10-mode) - 20°32'52 " (60-mode) Latitude north with an elevation of 22 meter above sea level near Gulf of Khambhat in the Arabian Sea, which is shown in the **figure 1**.

## Sampling of Water:

All the seventeen ground water samples were collected from manually operated wells and bores in polyethylene cans of 1L capacity separately for physico chemical analysis from 15<sup>th</sup> January 2012 to 17<sup>th</sup> January 2012. Before water sampling, all the polythene containers were cleaned and rinsed thoroughly with water samples to be analyzed. The physico-chemical and bio-chemical analysis were done using the standard methods.

## Experimental

All the seventeen water samples were examined for pH, Turbidity, Total Hardness, Total Alkalinity, Total Solid, Total Dissolved Solid, Total Suspended Solid, Nitrate, Phosphate, Fluoride, Chloride, Sulphate, BOD(Biological Oxygen Demand), COD(Chemical Oxygen Demand). All the chemical analysis were carried out using standard operating procedures with following methods and instruments or equipments shown in the Table 1.

A. Author works with Department of Chemical Engineering,  
Government Engineering College, Valsad( e-mail:  
[brsudani@yahoo.com](mailto:brsudani@yahoo.com))

restore the water quality better again. It therefore

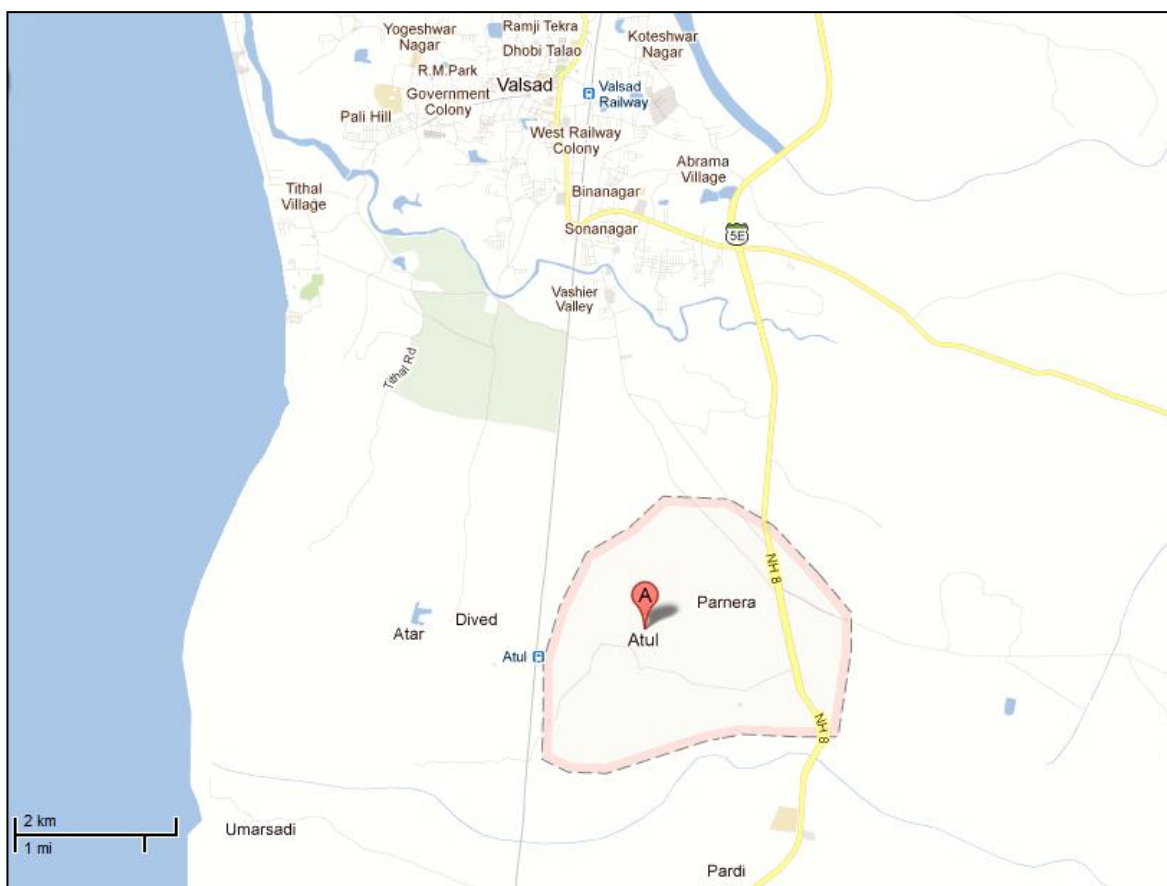


Fig. 1 Google Map of the study area.

**Table 1:** Analytical methods and equipment used in the study

Sr. No.	Physico-chemical Parameter	Method	Instruments/Equipment
1	pH	Electrometric	pH Meter
2	Turbidity	Nephelometric	Bench type Turbidimeter
3	Total Hardness	Titration by EDTA	-
4	Total Alkalinity	Titration by H <sub>2</sub> SO <sub>4</sub>	-
5	Total Solid	Gravimetrically	-
6	Total Dissolved Solid	Gravimetrically	-
7	Total Suspended Solid	Mathematically	TSS = TS - TDS
8	Nitrate	Ultraviolet screening	UV-VIS Spectrophotometer
9	Phosphate	Molybdophosphoric acid	UV-VIS Spectrophotometer
10	Fluoride	SPADNS method	UV-VIS Spectrophotometer
11	Chloride	Titration by AgNO <sub>3</sub>	-
12	Sulphate	Turbid metric	Turbidimeter
13	Biological Oxygen Demand	5 days incubation at 20° C followed by titration	BOD Incubator
14	Chemical Oxygen Demand	Digestion followed by titration	COD Digester
15	Total coliform	Multiple tube fermentation technique	Bacteriological Incubator
16	Faecal coliform	Multiple tube fermentation technique	Bacteriological Incubator

## II. Results and Discussion

All the seventeen water samples were analyzed for above physico-chemical and biochemical parameters. This hydro-chemical data was compared with reference to BIS [3] and

WHO [4] standards. The results of the chemical analysis of groundwater from this study are presented in Table-2A and Table-2B.

**Table 2A:** Results of Physico-chemical analysis.

Tests	pH	Trb	TH	TAL	TS	TDS	TSS	NT	PS	F	Cl	Sul
Sample												
S <sub>1</sub>	6.9	2.4	310	130	446	423	23	23	-	0.23	198	210
S <sub>2</sub>	6.8	3.2	290	122	425	400	25	14	-	0.34	230	167
S <sub>3</sub>	7.0	3.2	285	223	415	385	30	25	-	0.33	240	249
S <sub>4</sub>	7.2	1.9	455	234	638	623	15	21	-	1.12	233	263
S <sub>5</sub>	7.8	4.1	259	350	364	354	10	13	-	0.25	345	289
S <sub>6</sub>	7.6	3.8	279	285	394	382	12	14	-	0.74	210	300
S <sub>7</sub>	6.9	3.1	610	125	863	820	43	45	-	1.29	210	221
S <sub>8</sub>	7.5	2.9	515	310	736	705	31	34	-	0.75	278	287
S <sub>9</sub>	7.6	2.3	526	320	743	720	23	23	-	0.78	267	233
S <sub>10</sub>	8.0	4.2	321	409	474	440	34	34	-	0.43	370	310
S <sub>11</sub>	6.9	1.8	291	113	426	398	28	21	-	0.42	246	212
S <sub>12</sub>	7.4	3.2	312	276	451	421	30	29	-	0.83	240	267
S <sub>13</sub>	7.4	3.4	318	279	449	436	13	20	-	0.34	258	271
S <sub>14</sub>	7.5	2.3	298	322	430	409	21	39	-	0.78	260	287
S <sub>15</sub>	7.8	2.5	625	358	873	852	20	43	-	1.23	352	292
S <sub>16</sub>	7.4	2.2	457	299	645	626	19	20	-	0.87	254	208
S <sub>17</sub>	6.9	3.0	463	124	655	637	18	32	-	0.88	202	200

*Trb= Turbidity (NTU), TH=Total Hardness (ppm), TA= Total Acidity, TAL=Total Alkalinity(ppm), TS= Total Solid(ppm), TDS= Total Dissolved Solid(ppm), TSS=Total Suspended Solid(ppm), NT= Nitrate(ppm), PS= Phosphate(ppm), F=Fluoride(ppm), Cl=Chloride(ppm), Sul= Sulphate(ppm)*

All the groundwater samples from the study area had found colorless and odorless. But the Taste of the water showed some brackish water for some sampling areas.

The recorded pH values were well within the limit of 6.5 – 8.5. Turbidity of groundwater is always due to the presence of clay, microbes and partially contaminated organic compounds. In this work the turbidity of the samples were found from 1 to 5 NTU which is in the permuted limit. This water quality is due to the presence of calcium and magnesium salt. The values found lowest 285 ppm to highest 625 ppm in which some samples showing slightly higher than BIS and WHO limits. The causes of alkalinity in water are due to the presence of various basic inorganic or organic compounds. According to BIS it should be in the range of desirable limit 200 ppm to permissible limit 600 ppm. Results of present work were found within the limit. The palatability of water with a total dissolved

solids (TDS) level of less than about 600 ppm is generally considered to be good; drinking-water becomes significantly and increasingly unpalatable at TDS levels greater than about 1000 ppm [5]. Here result of analysis showed that all the values of TDS and TS were between the permissible limit.

The guideline value for nitrate of 50 mg/l was also found not violated. In setting national standards or local guidelines for fluoride, this experimental results were found in the recommended limit (<1.5 ppm). Chloride concentration in the form of chloride ion is one of the Major inorganic anion in water. All these water samples had chloride falling under permissible limits (200-1000 Mg/l).

The reported values of Sulphate amounts 167 ppm lowest to 310 highest were within the limit of 200-400mg/l at all the locations. In this study the amount of phosphate was found nil.

**Table 2B:** Results of Bio-chemical analysis.

Tests	BOD	COD	Total coliform	Faecal coliform
-------	-----	-----	----------------	-----------------

Sample	in ppm	in ppm	colonies per 100 ml	colonies per 100 ml
S <sub>1</sub>	1.12	2.01	8	4
S <sub>2</sub>	--	--	9	14
S <sub>3</sub>	0.50	--	10	10
S <sub>4</sub>	0.50	1.00	20	16
S <sub>5</sub>	--	--	11	12
S <sub>6</sub>	1.0	0.52	12	17
S <sub>7</sub>	--	--	2	4
S <sub>8</sub>	1.05	--	9	13
S <sub>9</sub>	--	--	4	5
S <sub>10</sub>	0.11	2.21	29	34
S <sub>11</sub>	0.21	--	9	12
S <sub>12</sub>	0.02	--	10	20
S <sub>13</sub>	--	--	20	13
S <sub>14</sub>	--	--	24	21
S <sub>15</sub>	1.00	1.00	21	12
S <sub>16</sub>	--	1.12	17	23
S <sub>17</sub>	--	--	12	9

Bio-chemical analysis showed that the amount of BOD were found nil for almost 47 % of the total samples. And the amount of COD were found nil for 65 % of the study samples. The values for Total coliform colonies were minimum 2 to maximum 29 per 100 ml and the values for faecal coliform colonies were minimum 4 to maximum 34 per 100 ml. It indicates that some minor contamination is present here in the higher valued samples.

### III. Conclusion and Recommendations

In respect of this study it is found that groundwater is the only source for people in the study area. The results of the chemical analyses of groundwater indicate considerable variation. The parameter PO<sub>4</sub><sup>3-</sup> is found as zero for all sites. Most of the water samples comply with WHO or BIS standards for drinking purposes. The higher proportions of some dissolved constituents are found in ground water, in some study area because of greater interaction of ground water with various materials in geologic strata or contamination. Although in most part the groundwater has not been severely polluted, from the groundwater conservation view of point, the groundwater still need protection and long term monitoring in case of future rapid industrial development.

### IV. Acknowledgements

Authors want to express their sincere appreciation to all the staff members of Chemical Engineering Department and Environment Engineering Department of Government Engineering College, Valsad for their helps and supports in

this analysis work. We are especially thankful to our Head of Department and Principal Sir for inspiring as well as helping in proof reading of this article. We heartily thankful to all the friends at different laboratories who helped us for the measurement and analysis work.

### References

- [1] Sinha Madhu Rani et al., "Physicochemical examination and quality assessment of groundwater (Hand-Pump) around Patna main town, Bihar state, India", *J. Chem. Pharm. Res.*, 3(3), 701-705 (2011)
- [2] Ramakrishnaiah C. R., Sadashivaiah C. and Ranganna G., "Assessment of Water Quality Index for the Groundwater in Tumkur Taluk, Karnataka State, India", *E-J Chem.*, 6(2), 523-530 (2009).
- [3] Bureau of Indian Standards, Indian Standards (IS: 10500) Drinking Water Specification: New Delhi (2004)
- [4] World Health Organization; Guidelines for drinking Water Quality: Vol.1, Recommendation 2<sup>nd</sup> Edition; Geneva, WHO (2008)
- [5] World Health Organization; Guidelines for drinking Water Quality: Recommendation 4<sup>th</sup> Edition; Geneva, WHO (2011)
- [6] Pandey Sandeep and Tiwari Sweta, "Physico-chemical analysis of ground water of selected area of ghazipur city-A case study", *Nature and Science*, 7(1), 17-20 (2009).
- [7] Bhaskar C.V. et al., "Assessment of Heavy Metals in Water Samples of Certain Locations Situated Around Tumkur, Karnataka, India", *E-J. Chem.* 7(2), 349- 352 (2010).
- [8] APHA, Standard Methods for Examination of Water and Waste Water, American Public Health Association, 20<sup>th</sup> Ed., Washington DC (1998)
- [9] Adekunle A.S., "Effects of Industrial Effluent on Quality of well water within Asa Dam Industrial Estate, Ilorin, Nigeria", *Nature and Science*, 7(1), 39-43 (2009).
- [10] LI PEI-YUE, QIAN HUI and WU JIAN-HUA, "Hydrochemical Characteristics and Evolution Laws of Drinking Groundwater in Pengyang County, Ningxia, Northwest China", *E-J Chem.*, 7(S1), S209-S216 (2010).



# Bio Based Plastics As Viable and Biodegradable Substitutes for Petroleum Based Counterparts

A. Neha Patni, B. Neha Tripathi

**ABSTRACT**—The problem of environmental pollution caused by indiscriminate dumping of plastic waste has assumed global proportions. These conventional plastics that are synthetically derived from petroleum are not readily biodegradable and are considered as environmentally harmful wastes. Its extensive use leads consequently to numerous environmental problems like pollution, recycling of waste plastics, and greenhouse effects. Various environmentally friendly substitutes of conventional plastics have been developed and are still being developed by chemical synthesis and microbial formulations. A material enabling not only to save oil consumption, but also to decrease CO<sub>2</sub> amount, is required to prevent global warming. The development of new techniques and materials that will enable to reduce CO<sub>2</sub> emission in the atmosphere has drawn the attention of many researchers in various fields. Bio-based plastics appear to be more environmentally friendly materials than their petroleum-based counterparts when their origin and biodegradability are compared. The paper reviews about the various substitutes of plastics like Cellulosic plastics, Poly lactides (PLA), Starch plastics, Polyhydroxyalkanoates (PHA), Liquidwood (bioplastics), WoodPowder Material, plastic from wheat gluten (WG) etc. which are nontoxic and biodegradable. Biodegradable plastics may serve as a promising solution to the overloaded landfills by diverting part of bulky volume plastics to other means of waste management, and to littering of disposable plastic products which are otherwise difficult to recycle.

**Index Terms**—biodegradable, Polyhydroxyalkanoates (PHA), Poly lactides (PLA), Wheat gluten (WG).

## I. INTRODUCTION:

**P**LASTIC materials are currently considered very important materials due to their exceptional properties and performance over other materials such as metal and wood. It is projected that the demand for plastics will continue to rise following a trend that has increased since the 1950s [1-3]. The annual world production of polymer materials was around 150 million tonnes in 1996, with the average per capita consumption of plastics in developed countries ranging from 80–100 kg per year.

A. Neha Patni:- Assistant Professor in Chemical Engineering Department, Nirma Institute of Technology, Nirma University, Ahmedabad, Gujarat, India. (Email id:- neha.patni @nirmauni.ac.in)

The rapid increase in production and consumption of plastics has led to the serious plastic waste problems, so called 'White Pollution' and landfill depletion, due to their high volume to weight ratio and resistance to degradation. Accumulated plastic film residues in soil have caused significant decrease in yield. Plastic wastes floating on rivers and lakes are increasingly threatening fishery, navigation, operation of hydropower plants, irrigation and other public works. Moreover, as over 99% of plastics are of fossil fuel origin, their rapid increase will put further pressure on the already limited non-renewable resources on earth [4].

Biodegradable plastics of renewable resources origin help to preserve the non-renewable resources and contribute to sustainable development. Some of the various biodegradable substitutes of plastics are Cellulosic plastics, Poly lactides (PLA), Starch plastics, Soy plastics, Liquid wood (bioplastics), etc. which are nontoxic, biodegradable, environmentally friendly. These various substitutes can be explained as:

## II. WHEAT GLUTEN

Among all the cereal and other plant proteins, wheat gluten (WG) is unique in its ability to form a cohesive blend with viscoelastic properties once plasticized. Generally, WG based plastics require plasticizing agents because of their poor toughness and water resistance. The intermolecular interactions and phase structures of plasticized wheat protein materials have been proposed recently. The hydrogen bonding interactions between plasticizer and protein and the soluble protein components are important factors in determining molecular motion in wheat protein materials. Both factors play key roles in the formation of a cohesive and continuous matrix in the materials to provide excellent mechanical properties [6]. Well-known plasticizers of WG include water, [7-12] glycerol, [11,12] and sorbitol, Saturated fatty acids, [10] diethanolamine, and triethanolamine [8] were also reported to plasticize WG. The use of a plasticizer normally enhances the elongation of WG but significantly reduces the strength. Blending with polymers (e.g., polycaprolactone (PCL), poly- (hydroxyesterether), maleic anhydride-modified polycaprolactone, [13] poly(ethylene-covinyl acetate) /poly(vinyl chloride), cassava starch, poly(butylene succinate) and poly(lactic acid) [12-14] is an alternative approach to improve the properties of WG; however, some of these polymers are too expensive, while others improve the mechanical strength at the expense of the loss of polymer elongation



AVOID				PREFER	
Polyvinyl Chloride (PVC)	Plastics with highly Hazardous Additives	Acrylonitrile Butadiene Styrene (ABS), Ethylene Vinyl Acetate (EVA), Polycarbonate, Polystyrene	Polyethylene (PE) Cross-linked (X)-PEX Polyethylene Terephthalate (PET)	Polyethylene Polypropylene Thermoplastic Polyolefin	Biobased plastics Sustainably grown

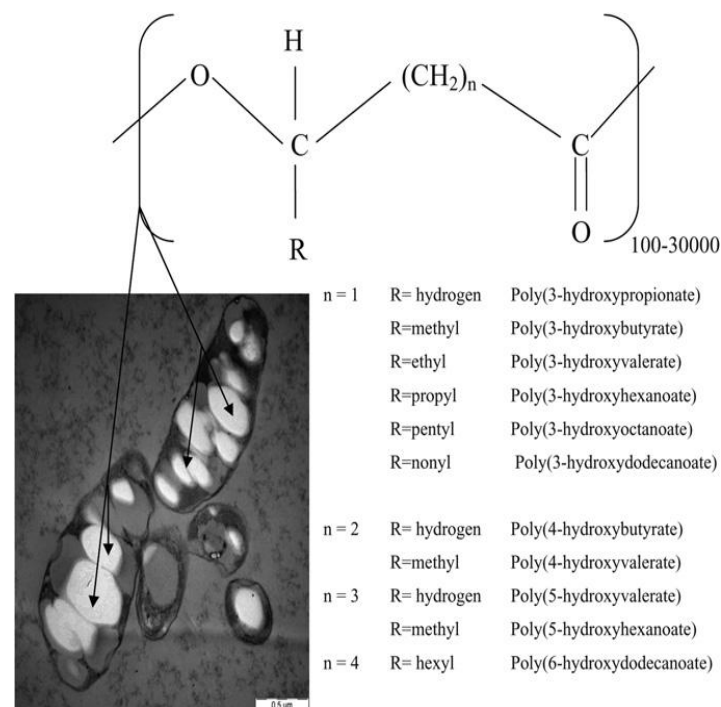
**Table 1:** Environmental preference spectrum for the health-care industry [5]

### III. POLYHYDROXYALKANOATES

The production of bioplastics from biorenewable sources such as waste effluents can make it more sustainable and can reduce our environmental footprint [14]. Polyhydroxyalkanoates (PHA) are microbially synthesized 100% biodegradable polymer [15]. The main advantage of PHAs over other types of biodegradable plastic is that they do not require special environmental conditions and can be rapidly degraded in both aerobic and anaerobic conditions, thereby solving the space problem in landfill. By converting waste stream into biodegradable plastics higher value product can be produced from low value source and at the same time higher COD value is converted into polymer [16]. In order to make the PHA production economically viable, many goals have to be addressed simultaneously. Economic biotechnological plastic production depends on highly productive microorganisms and on low-cost substrates. The nature of the substrate not only determines the PHA content but also its composition, which subsequently affects the final polymer properties. The use of these inexpensive carbon sources to produce PHAs could lead to significant economical advantages.

The structural diversity of PHA, its biodegradability, and its close analogy to plastics make them extremely desirable substitute for synthetic petrochemical-derived plastics. There is a great potential supply of raw material to manufacture bioplastics derived from waste stream. The process may involve the bioconversion of Palm Oil Mill Effluent (POME) to organic acids, recovery of the acids and finally the utilization of these acids for the biosynthesis of PHA. POME originates from two main processes i.e., the sterilization and clarification stages; as condensate and clarification sludge, respectively. Depending on the process of oil extraction and the properties of fresh fruit bunch (FFB), POME is made up of about 95-96% water, 0.6-0.7% oil and 4-5% total solids (mainly debris from the fruit). No chemicals are added during palm oil production process and thus it is non-toxic waste. Upon discharge from the mill, POME is in the form of highly concentrated dark brown colloidal slurry of water, oil, and fine cellulosic materials. POME is highly polluting and 100 times more polluting than domestic sewage in terms of BOD and COD, which generally exceed to 25,000 mg/L and 50,000mg/L, respectively.

Polyhydroxyalkanoates (PHAs) are accumulated as a carbon and/or energy storage material (Fig. 2) in various microorganisms usually under the condition of limiting nutritional elements such as N, P, S, O, or Mg in the presence of excess carbon source [17]. The treatment involves a two-step process whereby the wastewater is first anaerobically treated to produce acetic and propionic acids which is used in the second stage for PHA synthesis by *Rhodobacter sphaeroides*[18]. The diluted acid obtained during anaerobic treatment is concentrated and separated using ion exchange resin for further use in PHA fermentation.



**Fig. 1:** The general structure of polyhydroxyalkanoates and its representatives and transmission electron micrographs of PHBV granules in *Comamonas* [16]

### IV. WOOD POWDER MATERIAL

A plastic-like material is obtained by compression of wood powder under an appropriate temperature and pressure conditions. The static bending strength of the wood powder material (WPM) at 23 °C is roughly the same as that of plastics such as ABS, and due to the auto-condensation of wood components the WPM can be enhanced and hardened. Since woods are sustainable biomass materials as well as environment-friendly resources, industrial wood-based products promise a good future. Due to the photosynthesis ability of trees they can adsorb CO<sub>2</sub> in the atmosphere to make woods. Moreover, they can be processed by less energy methods such as sawing and planing, compared with those of metallic and polymer materials. During their uses as woods, there is no emission of CO<sub>2</sub>, and even during disposal stage, only carbon dioxide and water are generated, which are needed by trees in their photosynthesis activity. This means that longer period uses of woods result in less CO<sub>2</sub> emission to the atmosphere due to stock as wood components in their body. In wood forming such as compression and bending of lumber, it is essential for the deformation of wood with small pressure to adjust not only to temperature but also moisture conditions. In terms of WPM production, the moisture, however, would diminish during the compaction of wood powder due to evaporation, and there exists a limitation pressure for preparing sound WPM by the compaction under the atmosphere pressure condition related to the boiling temperature of water [19].

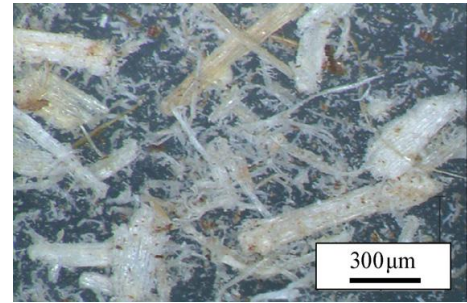
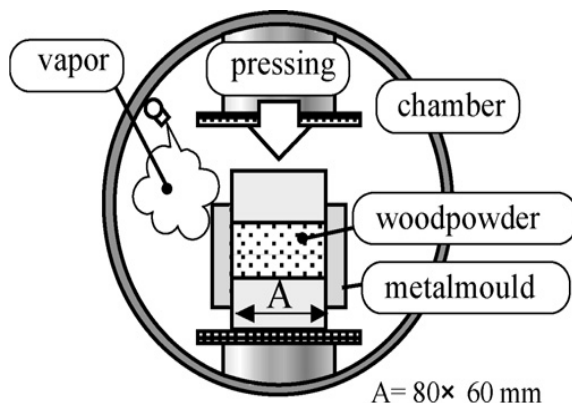


Fig. 2: Schematic Diagram of steaming pressing and microscopic WPM [19]

## V. LIQUID WOOD

Liquid wood is a promising new bioplastic, or biopolymer. These are the materials that look, feel and act just like plastic but unlike petroleum-based plastic, they're biodegradable. In this case, liquid wood is made from pulp-based lignin, a renewable resource. To make biopolymers, lignin, a by product of paper mills, is mixed with water and exposed to high temperatures and high pressure to create a mouldable composite material that's strong and nontoxic, making it a good plastic substitute. German researchers have used it to manufacture a variety of items including toys, golf tees, speaker boxes, etc. Since made of wood, it can be recycled as wood, either broken into pieces and used as filler or burned.

## VI. CELLULOSIC PLASTICS

Cellulosic and lignin plastics are produced by chemical modification of natural cellulose and lignin obtained from wood and short cotton fibres called linter. Natural fibres are typically added as reinforcement to plastics [20]. Modified natural polymers such as cellulose acetate have been used instead of conventional plastics, for various purposes. Often these cellulose acetate films resemble plastics in their hydrophobic nature and strength. The high levels of acetylation required to modify natural polymers such as cellulose, starch, and xylan for use as substitute plastics have been demonstrated to reduce the anaerobic biodegradation potential dramatically.

## VII. POLYLACTIDES

PLA is a thermoplastic aliphatic polyester obtained by the polymerization of lactic acid derived from microbial fermentation of corn starch or cane sugar. Pure thermoplastic starch (TPS) is obtained without fermentation or chemical treatment of natural corn, potato, rice, tapioca or wheat starch that are extruded or blended to produce TPS. They decompose within 47 days, without emitting toxic fumes when burned and manufacturing them uses 20 to 50 percent less fossil fuels than petroleum-based plastics. Corn-based plastics can be used to manufacture food storage containers as well as storage for consumer goods.[21-22]

## VIII. COMPARISON

**Table 2.** Sustainability improvements of bio-based plastics relative to petroleum-based plastics (PBP) [23]:

Bioplastic	Sustainability improvement
Polyhydroxyalkanoates (PHAs)	Highly biodegradable.
Polylactic acid (PLA)	Production uses 30-50% less fossil energy and generates 50-70% less CO <sub>2</sub> emissions than PBP. Competitive use of water with the best performing PBP, recyclable, compostable at temperatures above 60 deg Celsius.
Cellulose and Lignin	The biological degradation of lignin is lower than cellulose, compostable.
Corn zein and soy protein	Biodegradable and compostable.

## IX. SUMMARY AND FUTURE PERSPECTIVES

Wheat Gluten can fully biodegrade after 36 days in aerobic fermentation and within 50 days in farmland soil without releasing toxic products which make it an ideal candidate for development of biodegradable substitute of plastic. Polyhydroxyalkanoates (PHAs) are thermoplastics, having similar mechanical properties to those of polypropylene, with the additional advantage of being completely biodegradable, biocompatible, and produced from renewable resources (sugars and fatty acids). PLA is carbon-neutral because it is derived from renewable plant starches. PLA is often made from genetically modified corn, the growing of which raises environmental, human-health and economic concerns, and the end result may take longer than estimated to decompose.

In spite of large number of excellent studies dedicated to characterising structure, important gaps of knowledge has to be filled by future research. One of priorities should be to get more insight to the changes of structure. A little understanding of the relation between structure and functionality of the polymer, particularly in combination with additive is essential. Plasticizers usually play important role in polymer compounding, so various possible plasticizer need to be explored. Types of plasticizers and delivery methods and better plasticizer combinations for biodegradable substitutes need to be identified.

## REFERENCES

[1] Azapagic, A., Emsley, A., Hamerton, I., 2003. *Polymers: The Environmental and Sustainable Development*. John Wiley and Sons, England, pp. 219.

[2] PlasticsEurope, EUPC, EPRO, EuPr, 2008. *The Compelling Facts about Plastics. An Analysis of Plastics Production, Demand and Recovery for 2006 in Europe*.

[3] Rosato, D.V., Rosato, D.V., 2003. *Plastics Engineered Product Design*. Elsevier, USA, pp. 1.2, 20-22, 588.

[4] X. Ren, Biodegradable plastics: a solution or a challenge? , *Journal of Cleaner Production* 11 (2003) 27–40

[5] Rossi, M., Lent, T. Creating Safe and Healthy Spaces : Selecting Materials that Support Healing The Center for Health Design, U.S.A. 2006

[6] Zhang, X.; Burgar, I.; Do, M. D.; Lourbakos, E. *Biomacromolecules* 2005, 6, 1661–1671.

[7] Marion Pommet, Morel et al ; Intrinsic influence of various plasticizers on functional properties and reactivity of Wheat gluten thermoplastic materials; 2005

[8] Narendra Reddy, Yiqi Yang ; Bio composites developed using water-plasticized WG matrix and jute bars as reinforcement ; *Polymer International* (2011); pi.3014

[9] H  l  ne Angellier-Coussy ,Emmanuelle Gastaldi,Nathalie Gontard, Val  rie Guillard ; Influence of processing temperature on the water vapour transport properties of WG based agromaterials *Industrial Crops and Products* Vol 33, Issue 2, 2011, Pages 457-461

[10] Sudsiri Hemsri,Alexandru D. Asandei, Kasia Grieco, Richard S. Parnas Biopolymer composites of WG with silica and alumina, *Composites Part A: Applied Science and Manufacturing* Vol 42, Issue 11, November 2011, Pages 1764-1773

[11] Lidia S Z  rate-Ram  rez, Inmaculada Mart  ne, Alberto Romero, Pedro Partal, Antonio Guerrero; Gluten-based materials plasticised with glycerol and water by thermoplastic mixing and thermomoulding; 2010

[12] Qiaolong Yuan , Wubin Lu, Yongkang Pan Structure and properties of biodegradable wheatgluten/attapulgitite nanocomposite sheets *Polymer Degradation and Stability* Volume 95, Issue 9, 2010, Pages 1581-1587

[13] Marion Pommet, Andr  as Redl, St  phane Guilbert, Marie-H  l  ne Morel; More Intrinsic influence of various plasticizers on functional properties and reactivity of WG thermoplastic materials *Journal of Cereal Science* Vol 42, Issue 1, 2005, Pages 81-91

[14] Khanna, S., Srivastava, A.K., 2005. Recent advances in microbial polyhydroxyalkanoates. *Process Biochem.* 40 (2), 607-619.

[15] Dovi, V.G., Friedler, F., Huisingh, D., Klemes, J.J., 2009. Cleaner energy for sustainable future. *J. Cleaner Prod.* 17 (10), 889-895.

[16] Tabassum Mumtaz, Noor Amalina Yahaya et al. Turning waste to wealth-biodegradable plastics polyhydroxyalkanoates from palm oil mill effluent : a Malaysian perspective, *Journal of Cleaner Production* 18 (2010) 1393-1402.

[17] Anderson, A.J., Dawes, E.A., 1990. Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol. Rev.* 54, 450-4

[18] Hassan, M.A., 2006. A process for treatment of palm oil mill effluent and for conversion of the palm oil mill effluent into biodegradable plastics. Malaysian Patent 123658- A.

[19] T. Miki, K. Takeuchi, H. Sugimoto, K. Kanayama Performance Study of compact wood powder material processing for improved impact characteristics aiming at substitute for plastics: *Journal of Materials Processing Technology* 192–193 (2007) 422–427

[20] Mohanty, A.K., Misra, M., Hinrichsen, G., Biofibres, biodegradable polymers and biocomposites: an overview. *Macromolecular Materials and Engineering* 276-277 (1) 2000., 1-24.

[21] Vink, E. Responsible Innovation: Reducing the Environmental Footprint with NatureWorks\_ *Biopolymer*. 2008

[22] Crank, M., Patel, M., Marscheider-Weidemann, F., Schleich, J., H  sing, B., Angerer, G., 2005. Techno-Economic Feasibility of Large-Scale Production of Bio-Based Polymers in Europe.

[23] Clara Rosal  a Alvarez-Chavez, Sally Edwards , Rafael Moure-Eraso , Kenneth Geiser, Sustainability of bio-based plastics: general comparative analysis and recommendations for improvement, *Journal of Cleaner Production* 23 (2012) 47-56.

# Exploration of New Technique for Removal of Metal Artifacts in Dental CT Images

A. Hetal K. Mehta, B. Himanshu R. Dave

**ABSTRACT-** Presence of artifacts in images can create a big trouble, especially in medical field. Due to artifacts, diagnosis cannot be done sufficiently and it will create problem for further application. Streak artifacts appear on maxillofacial X-ray computed tomography (CT) images. It arises due to presence of any metal implant in the oral cavity. So it causes problem when 3d model is generated from it. In this paper, we propose a method for removing metal streak artifacts present in dental ct scan images. This method includes fuzzy theory and matlab for MAR (metal artifact removal).

**Index Terms** – artifact, CT, fuzzy theory, MAR

## I. INTRODUCTION

Basically, the CT scan images represent cross-sections through the body. Projection data are collected and reconstructed into images as the patient is moved incrementally through the CT gantry. The 3D representation of Computed Tomography (CT) scans is widely used in medical applications such as virtual endoscopy, plastic reconstructive surgery, and also in dental implant planning systems and more. Metallic objects present in CT scans cause noticeable artifacts such as beam hardening and streaking, as shown in Fig. 1.

During CT scan, when x rays are impact to metal object such as dental restorations, surgical plates and pins and radiographic markers can cause this type of the artifact. Since metal having high attenuation coefficient it absorbs x rays and act as a source of artifacts.

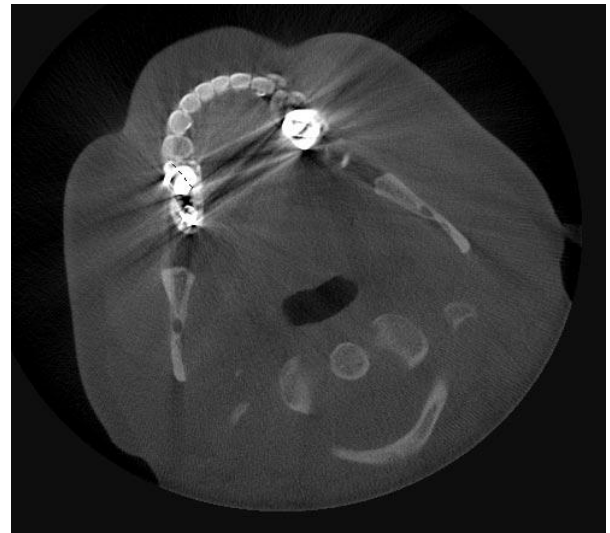


Fig 1:- Typical dental slice (real patient) with massive metal parts (crowns) within field of view.

When image is reconstructed, metal causes the brighter and darker streak artifacts. So image quality is degraded.



Fig 2:- Incorrect RP jaw model.

At result, important information is destroyed both at near dental fillings as well as in the periphery of the image because X-ray beams are completely absorbed by metallic implants.

A. PG Student at Government Engineering College, Gandhinagar, Gujarat-382028, India(Email: hetalmehtahkm@yahoo.in)

B. Professor in Electrical Engineering Department at Government Engineering College, Gandhinagar, Gujarat-382028, India (Email: hrd@gecg28.ac.in)

In surgical planning, detailed CT-based knowledge of the exact anatomy of the patient is fundamental, especially in the dental region. In order to reduce surgery time and pre-surgical planning, Rapid Prototyping (RP) systems are used to manufacture 3D anatomical replicas.

However, it is useless because it will generate 3D spikes in RP model as seen in fig 2. So metal artifacts are must be removed before model manufacturing.

Our goal is to localize artifacts from the main image and remove it. At last the values on those areas are replaced by linear interpolation.

## II. LITERATURE SURVEY

Recently, research is done in Department of Computer Science, Faculty of Engineering, Kitami Institute of Technology, 165, Koen-cho, Kitami, Hokkaido 090-8507, in Japan, on Successive iterative restoration applied to streak artifact reduction in X-ray CT image of dento-alveolar region. X-ray computed tomography (CT) images in the dento-alveolar region are sometimes rendered unusable for diagnostic purposes due to the appearance of streak artifacts. The purpose of the study is to reduce streak artifacts appeared on dental and maxillofacial X-ray CT images by the application of modified iterative restoration method. They took an advantage of the aspect that adjacent CT images often depict very similar anatomical structures within the resulting collection of thin-slice images.

CT images having streak artifacts were processed using the projection data of adjacent CT images. A modified iterative correction, the maximum likelihood-expectation maximization (ML-EM) reconstruction algorithm, was employed to reduce the streak artifact caused by metallic materials in the oral cavity. It approximates between the processed image and the original projection data. First, the projection data of an intact image were obtained, and then, the next image that had streak artifacts was processed. The projection data of the processed image were obtained, and the ML-EM method was applied to the next image again. Then, the successive iterative restoration was carried out [2].

Other studies also support the conclusion that the Multichannel CT allows faster scanning times, resulting in reduced motion artifacts; thinner sections, with which it is possible to create a scanned volume of isotropic voxels with equivalent image resolution in all planes; and the generation of a higher X-ray tube current, which may result in better penetration of metal hardware and reduction of artifacts [3].

similarly, research conclude that An integrated and effective metal artifact reduction method named Metal Erasing (ME) especially suited to dental applications is proposed.[4] Layout of metals is identified as metal-only tomogram, using its characteristics of X-ray opacity and

simple image processing technique of linearization together with backward projection. Metal-only sinogram is calculated by forward projection of the metal-only tomogram, and identifies corrupted areas on the original sinogram. The areas are then replaced by interpolation, and filtered back projection (FBP) produces a tomogram without figures of metals.

In 2004, research has done on a pragmatic approach to metal artifact reduction in CT: merging of metal artifact reduced images by Watzke O, Kalender WA [5]. The purpose of this study was to improve metal artifact reduction (MAR) in X-ray computed tomography (CT) by the combination of two artifact reduction methods.

The presented method constitutes an image-based weighted superposition of images processed with two known methods for MAR: linear interpolation of reprojected metal traces (LI) and multi-dimensional adaptive filtering of the raw data (MAF). Two weighting concepts were realized that take into account mean distances of image points from metal objects or additional directional components. Artifact reduction on patient data from the jaw and the hip region shows that although the application of only one of the MAR algorithms can already improve image quality, these methods have specific drawbacks. While MAF does not correct corrupted CT values, LI often introduces secondary artifacts. The corrective impact of the merging algorithm is almost always superior to the application of only one of the methods. The results obtained with directional weighting are equal to or in many cases better than those of the distance weighting scheme. Merging combines the advantages of two fundamentally different approaches to artifact reduction and can improve the quality of images that are affected by metal artifacts.

Study on “reduction of CT artifacts caused by metallic implant” introduces a technique to reduce metallic artifacts present in ct image. The missing project data are replaced by linear interpolation[7].

## III. APPROACH FOR METALLIC IMPLANT ARTIFACTS REMOVAL

The proposed method consists of following two Steps: 1. Clustering and 2. Interpolation.

Clustering means to divide similar data into same groups. Clustering methods that can be used to organize data into groups based on similarities among the individual data items. Here, fuzzy clustering is used for identify the artifacts.

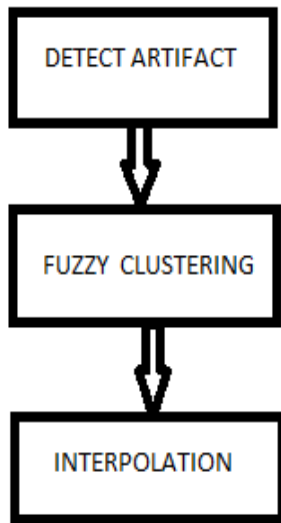


Fig 3:-approach for metallic implant artifact removal

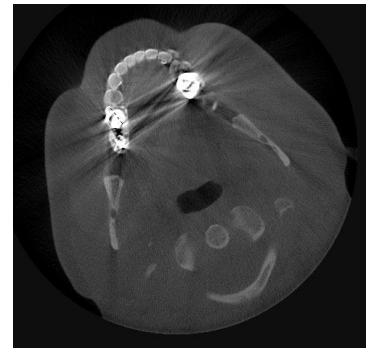
These algorithms are used to separate the regions affected by artifacts from the regions where the sources are localized and other parts of image. Fuzzy logic is having lots of clustering application over simple method of clustering.

In this research, we have applied fuzzy K mean clustering which is a more statistically formalized method and discovers soft clusters where a particular point can belong to more than one cluster with certain probability. Interpolation is the process of determining the values of a function at positions lying between its samples. It achieves this process by fitting a continuous function through the discrete input samples. This permits input values to be evaluated at arbitrary positions in the input, not just those defined at the sample points.

The interpolation techniques are divided into two categories, deterministic and statistical interpolation techniques. Here we apply linear interpolation technique for generating missing data in image.

IV. RESULTS

Figure 1 shows the main image having metal streak artifacts on which this experiment has done. Firstly clustering i.e. k-mean clustering is applied on image. Data of image is segmented into three clusters showing three different parts of the information. Results of clustering techniques are as shown in figure 4.



(a)



(b)



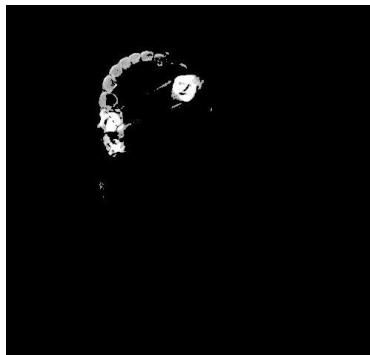
(c)



(d)

Fig 4:- results of fuzzy k-mean clustering.(a) Original image of dental CT. (b), (c), (d) are showing three different portion of main image.

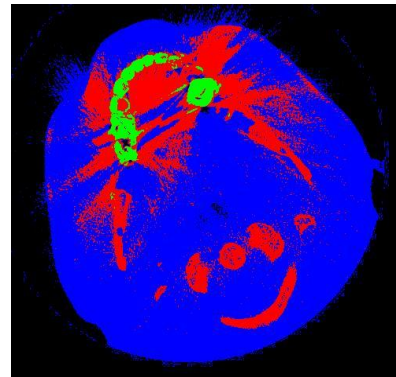
With using matlab and image processing toolbox different more logical operation is done, results of it is as shown in figure 5.



(a)



(b)



(d)

Fig:-5 (a) only teeth without artifacts. (b) Whole image without teeth and (c) color is filled in 3 clusters.

In figure 5(C) gives the clear idea of three different regions from that we get the normal image i.e image which is not having artifacts. At last interpolation is done by using interpolation algorithm in matlab. As a result, we get better image for diagnosis purpose. Figure 6 shows the output image.

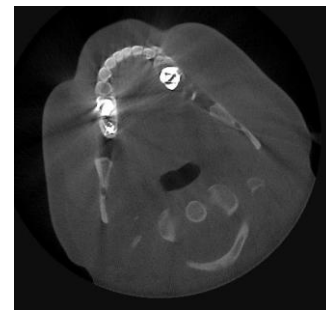


Fig 6:-output image, without streak artifacts.

### V. CONCLUSION

This method is capable for reducing metal artifacts in regions where metallic objects are present. It is observed that this method gives sufficient output for diagnosis purpose or any further applicaions.3D model reconstruction can also be done correctly so that doctors can know the exact anatomy of patient.

With use of fuzzy theory, we get eliminated artifact trails, without damage to the visualization of metallic presences in the image, and thus allowing applying the interpolation algorithms directly on the image.

### REFERENCES

- [1] Massimo Martorelli,(2011),A Novel Method of Removing Artifacts in Dental CT Images, Proceedings of the IMProVe 2011 International conference on Innovative Methods in Product Design June 15<sup>th</sup> – 17th, 2011, Venice, Italy
- [2] Atsushi Kondo A., Hayakawa Y., Dong J., Honda A. (2010), Iterative correction applied to streak artifact reduction in an X-ray computed tomography image of the dento-alveolar region, Oral Radiol 26:61–65 DOI 10.1007/s11282-010-0037-6.
- [3] Stradiotti P., Curti A., Castellazzi G., Zerbi A. (2009), Metal-related artifacts in instrumented spine. Techniques for reducing artifacts in CT and MRI: state of the art, Euro Spine Journal, S 102-S108.
- [4] A Practical Method to Reducing Metal Artifact for Dental CT Scanners (2008), kobayashi-koji@jp.yamatake.com
- [5] Barrett JF, Keat N (2004) Artifacts in CT: recognition and avoidance. Radio Graphics 24:1679–1691.
- [6] Watzke O., Kallender W. (2004), A pragmatic approach to metal artifact reduction in CT: merging of metal artifact reduced images, European Radiology 14, 849– 856.
- [7] Kalender W., Hebel R., Ebersberger J. (1987), Reduction of CT artifacts caused by metallic implants, Radiology 164, 576–577.
- [8] Reduction of artifacts due to multiple metallic objects in computed tomography. Kye Young Jeong and Jong Beom Ra, Dept. of EECS, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, [www-isl.kaist.ac.kr](http://www-isl.kaist.ac.kr)
- [9] Medical decision making through fuzzy computational intelligent approaches , Department of Informatics and Computer Technology, Technological Educational Institute (TEI) of Lamia, 3rd km Old National Road Lamia-Athens, 35100 Lamia, Greece,Elpiniki I. Papageorgiou, [epapageorgiou@teilam.gr](mailto:epapageorgiou@teilam.gr)
- [10] Oral and maxillofacial radiology , Michel W. Vannier, MD, Charles F. Hildebolt, DDS, PhD, university of Iowa and mallinckrodt institute of radiology, washington university school of medicine, November 1977.



# Theoretical calculation of electron impact total ionization cross sections for aluminum atom over a range of energies (Ionization threshold – 2000 eV)

A Chetan Limbachia, B Ashok Chaudhari, C Harshad Bhutadia

**Abstract** - Total ionization cross sections,  $Q_{ion}$ , for electron impact on aluminium (Al) and is reported over a wide energy range from ionization threshold to 2 keV. Spherical Complex Optical Potential (SCOP) formalism [1, 2] is employed to calculate electron impact total elastic and inelastic cross sections. To extract total ionization cross sections from calculated total inelastic cross sections, we employ CSP-ic (Complex spherical potential – ionization contribution) formalism [1, 2]. Our calculated data for Al has found overall good agreement with the available theoretical and experimental data.

**Index Terms**: Electron scattering, Ionization cross section, Inelastic cross section.

## I. INTRODUCTION

ELECTRON impact collision data are in ever increasing demand for many decades. This is the consequence of their utility in various fields of applied physics. Hence considerable progress has been made on electron atom/molecule collision studies both theoretically as well as experimentally.

Electron-impact ionization is one of the most fundamental collision processes in atomic and molecular physics. It sustains gas discharges and plasmas, leads to most of the chemistry in radiation effects, plays a major role in planetary upper atmospheres, and is the basis for much of mass spectrometry. Because of their basic and practical importance, cross sections for electron-impact ionization have been measured since the earliest days of atomic collision physics [3, 4].

In this paper we present electron impact total ionization cross sections,  $Q_{ion}$  for Al over a impact energies from ionization threshold to 2 keV. Calculated ionization cross sections are listed in table 1 and shown in graph 1. Our calculated data has overall good agreement with previous experimental and theoretical results (wherever available).

## II. THEORETICAL METHODOLOGY

The electron atom/molecule scattering phenomenon is characterized quantitatively by two important cross sections viz. total elastic and total inelastic cross sections and they combine to represent total cross sections. Accordingly we have,

$$Q_T(E_i) = Q_{el}(E_i) + Q_{inel}(E_i) \quad (1)$$

where the first term on the right hand side accounts for all elastic processes while the second term takes care of loss of flux in the outgoing channel resulting from electronic excitations and ionization. We are interested here in the total ionization cross sections and hence the second term of equation (1) is important. The inelastic processes are taken into account through the complex part of the optical potential via absorption potential. The complete spherical complex optical potential (SCOP) [5] is represented by

$$V_{opt}(E_i, r) = V_R(E_i, r) + iV_I(E_i, r) \quad (2)$$

where, the real part  $V_R$  consists of static potential ( $V_{st}$ ), exchange potential ( $V_{ex}$ ), and polarization potential ( $V_p$ ). Owing to the fixed nuclei approximation, the static potential, ( $V_{st}$ ) is calculated at the Hartree-Fock level. The exchange potential ( $V_{ex}$ ) is responsible for electron exchange between the incoming projectile and the target-electrons. The polarization potential ( $V_p$ ) combines the short range correlation and long range polarization effect that arises due to the momentary redistribution of target charge cloud which gives rise to dipole and

A P. S. Science College, Kadi 382 715, Gujarat, India  
 B Government Engineering College; Patan-384 265, Gujarat, India  
 (E mail: akc12878@yahoo.com)  
 C Government Engineering College; Patan-384 265, Gujarat, India  
 (Mobile No. 09376048550, E mail: harshbhutadia@gmail.com)

quadrupole moments. The second term of equation (2) is the imaginary part of the potential which is taken care by absorption potential. It is to be noted here that the spherical complex optical potential (SCOP) as such does not require any fitting parameters. All the potentials described vide equation (2) are charge-density dependent. Hence, representation of target charge density is very crucial. We have employed atomic charge density derived from the Hartree Fock wave functions of Bunge and Barrientos [6]. In order to represent the target molecule we have adopted single centre approach where the charge-density of all constituent atoms is expanded at the centre of mass of the system. The molecular charge density,  $\rho(r)$ , so obtained is renormalized to incorporate the covalent bonding as described in our earlier paper [7]. In the SCOP [7] formalism, the spherical part of the complex optical potential is used to solve exactly the Schrödinger equation using partial wave analysis to yield various cross sections [5]. Presently our absorption potential is elastic to both vibrational and rotational excitations of the target.

As discussed earlier the absorption potential takes care of loss of flux into all allowed inelastic channels. For this we have used model potential of Staszewska *et al* [8] which is non empirical, quasifree, Pauli-blocking and dynamic in nature. The full form of model potential is represented by

$$V_{abs}(r, E_i) = -\rho(r) \sqrt{\frac{T_{loc}}{2}} \times \left( \frac{8\pi}{10k_F^3 E_i} \right) \times \theta(p^2 - k_F^2 - 2\Delta) \cdot (A_1 + A_2 + A_3) \quad (3)$$

The parameters  $A_1$ ,  $A_2$  and  $A_3$  are defined as,

$$A_1 = 5 \frac{k_f^3}{2\Delta}; \quad A_2 = \frac{k_f^3 (5p^2 - 3k_f^2)}{(p^2 - k_f^2)^2}; \quad (4)$$

$$A_3 = \frac{2\theta(2k_f^2 + 2\Delta - p^2)(2k_f^2 + 2\Delta - p^2)^{5/2}}{(p^2 - k_f^2)^2}.$$

The local kinetic energy of the incident electron is given by

$$T_{loc} = E_i - (V_{st} + V_{ex}) \quad (5)$$

The dynamic absorption potential is density functional wherein it depends on charge density ( $\rho(r)$ ) of the target, incident energy ( $E_i$ ) and the parameter  $\Delta$  of the target. It is sensitive to short range potential like static and exchange through term  $T_{loc}$  and insensitive to long range potentials like polarization. In equation (3),  $p^2 = 2E_i$ , represents the momentum transfer of incident electron in Hartree, and  $k_F = [3\pi^2 \rho(r)]^{1/3}$  is the Fermi wave vector.

Also,  $\theta(x)$  is the Heaviside unit step-function which depends on  $p$ ,  $k_F$  and  $\Delta$ , such that  $\theta(x) = 1$  for  $x \geq 0$ , and is zero otherwise. In other words  $\theta(x)$  defines the boundary below which the absorption potential is zero and above which it has finite value. The dynamic parameters  $A_1$ ,  $A_2$  and  $A_3$  of equation (3) are the functions of  $\rho(r)$ ,  $I$ ,  $\Delta$  and  $E_i$ .

The parameter  $\Delta$  is very important since it determines a threshold below which  $V_{abs} = 0$ , implying that the ionization or excitation channels are prevented energetically. This further infers that the  $\Delta$  parameter represents the threshold energy for continuum states, which means only ionization process is taken into account, excitation to discrete levels being ignored by the original model [9]. So in order to include the excitations due to discrete levels at lower energy, we have considered  $\Delta$  as the energy dependent parameter. A variable  $\Delta$  accounts for more penetration of the absorption potential in the target charge-cloud region [1, 2]. Following the earlier works in this regard [1, 2], we express  $\Delta$  as a function of  $E_i$  around  $I$  as

$$\Delta(E_i) = 0.8I + \beta(E_i - I) \quad (6)$$

Here,  $\beta$  is obtained by requiring that  $\Delta = I$  at  $E_i = E_p$ , where  $E_p$  is the value of  $E_i$  at which  $Q_{inel}$  attains maximum value. For  $E_i > E_p$ ,  $\Delta$  is held constant equal to Ionization energy of the target as suggested in the original model of Staszewska *et al* [8].

After generating the full complex optical potential given in equation (2) for a given electron molecule system, we solve the Schrödinger equation numerically with Numerov method using partial wave analysis. At low incident electron energies with short range potentials, only few partial waves are significant for convergence, e.g. at ionization

threshold of the target around 5-6 partial waves are sufficient but as the incident energy increases large number of partial waves are needed for convergence. Using these partial waves the complex phase shifts are obtained which are key ingredients to find the relevant cross sections. The phase shifts contains all the information regarding the scattering event.

Total inelastic cross section determined vide equation (1) is not a directly measurable quantity and hence not directly comparable quantity. But, it is one of the most important quantities as it contains the

ionization and electronic excitations which are directly measurable quantities. Thus, we partition the total inelastic cross sections into its two vital components one due to the discrete electronic excitations and other due to the continuum ionization contribution, as,

$$Q_{inel}(E_i) = \sum Q_{exc}(E_i) + Q_{ion}(E_i) \quad (7)$$

Here, first term represents the sum over total excitation cross sections for all accessible electronic discrete transitions, while the second term is the total cross section due to all allowed electronic transitions to continuum i.e. ionization. In the present range of energies it is the single ionization that dominates in equation (7). The discrete transitions arise mainly from the low-lying dipole allowed transitions for which the cross section decreases beyond  $E_p$ . By definition,

$$Q_{inel}(E_i) \geq Q_{ion}(E_i) \quad (8)$$

Total ionization cross section may be estimated from total inelastic cross section by defining an energy dependent ratio  $R(E_i)$  given by

$$R(E_i) = \frac{Q_{ion}(E_i)}{Q_{inel}(E_i)} \quad (9)$$

such that,  $0 < R \lesssim 1$ .

As total ionization cross section is a continuous function of energy, we can express this ratio also as a continuous function of energy for  $E_i > I$ , used in earlier studies as [1, 2]

$$R(E_i) = 1 - f(U) = 1 - C_1 \left( \frac{C_2}{U+a} + \frac{\ln(U)}{U} \right) \quad (10)$$

here,  $U$  is the dimensionless variable defined by,

$$U = \frac{E_i}{I}$$

The reason for adopting such explicit form of  $f(U)$  could be visualized as follows. At high energies the total inelastic cross section follows the Born Bethe term according to which the cross sections falls of as  $\ln(U)/U$ , but at low and intermediate energies they obey  $1/E$  form [10]. Accordingly the first term will take care of the cross section behavior at low and intermediate energies while the second term will take care at high energy. The dimensionless parameters  $C_1$ ,  $C_2$  and 'a' involved in the above equation are deduced by imposing the three conditions on the ratio as discussed below.

$$R(E_i) \begin{cases} = 0 & \text{for } E_i \leq I \\ = R_p & \text{for } E_i = E_p \\ \cong 1 & \text{for } E_i \gg E_p \end{cases} \quad (11)$$

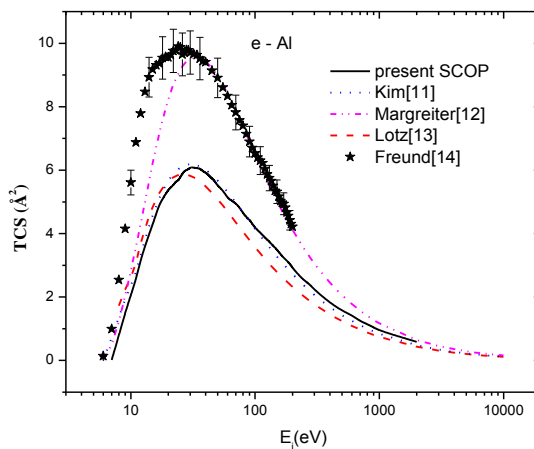
The first condition is an exact condition wherein it states that no ionization process is possible below the ionization threshold of the target implying that the value of the ratio must be zero. Coming to the last condition, which physically states that ionization contribution is almost equal to inelastic contribution at very high ( $\sim 10 E_p$ ) energies, this is attributed to the fact that at such high energies there are innumerable channels open for the ionization as against very few finite channels for excitation. At such high energies the contribution of excitation is almost negligible. Thus the ratio approaches unity.

The second condition is very crucial and was empirical in nature in CSP-ic method.  $R_p$  is the value of  $R$  at  $E_i = E_p$ , and it was observed that at the peak of inelastic cross section the contribution for ionization is about 70 to 80 %. And this argument

was supported by many targets studied through CSP-ic[1 - 2].

### III RESULTS AND DISCUSSION

The theoretical approach of SCOP along with the CSP-ic method outlined above is employed to determine total inelastic cross sections,  $Q_{inel}$  and total ionization cross sections,  $Q_{ion}$  along with a useful byproduct of electronic excitations in terms of the summed cross section  $\sum Q_{exc}$ . In the present paper, we have investigated aluminum atom of our interest and computed the total ionization cross sections using the CSP-ic method.



**Figure 1: Total ionization cross sections for e – Al scattering in  $\text{Å}^2$ .**

Solid line → Present SCOP, Dot line → Kim and Stone [11], Dash dot dot line → Margreiter *et al* [12], Dash line → Lotz [13], Star → Freund *et al* [14].

The ionization cross sections are plotted as function of projectile energy for Al, from threshold of the target to 2 keV as shown in figure 1. The numerical values of the total ionization cross section are tabulated in Table 1.

In figure 1, we show the comparison of present total ionization cross sections for e – Al scattering with available data. There is only one measurement reported by Freund *et al* [13] with experimental uncertainty of 7%. There are three theoretical

calculations available for e – Al scattering, Kim and Stone [11], Margreiter *et al* [12], Lotz [13]. Our calculated electron impact ionization cross section data for e – Al scattering by SCOP and CSP – ic, have overall good agreement with the data of Kim[11] and Lotz[13].

Table 1: Total ionization cross section (TICS) for Al

Energy (eV)	TICS ( $\text{Å}^2$ )
10	2.07
15	4.29
20	5.37
30	6.08
40	5.93
50	5.51
60	5.12
70	4.95
80	4.65
90	4.43
100	4.18
200	2.92
300	2.23
400	1.82
500	1.57
600	1.43
700	1.24
800	1.14
900	1.04
1000	0.94
2000	0.59

### IV CONCLUSION

We can conclude from our study based on SCOP and CSP – ic method, that our data found overall good agreement with available data. Data calculated by such reliable method can use to create a database for application point of view.

### V ACKNOWLEDGEMENT

CGL thank UGC, New Delhi for major research project under which part of this work is done. AKC and HSB are thankful to authority of government engineering college, Patan, for their kind support towards the research.

## VI REFERENCES

- [1] M. Vinodkumar, K. Korot, H. Bhutadia, Int. J. Mass. Spectrom, **294** (2010) 54.
- [2] M. Vinodkumar, R. Dave, H. Bhutadia, B. K. Antony, Int. J. Mass. Spectrom, **292** (2010) 7–13.
- [3] P. T. Smith, Phys. Rev. **36**, 1293 (1930).
- [4] W. Bleakney, Phys. Rev. **34**, 157 (1929); **35**, 139, 1180 (1930); **36**, 1303 (1930); **37**, 808 (1931).
- [5] A. Jain, K. L. Baluja, Phys. Rev. A **45(1)** (1992) 202.
- [6] C. F. Bunge, J. A. Barrientos, At. Data Nucl. Data Tables, **53** (1993) 113.
- [7] M. Vinodkumar, K. N. Joshipura, N. J. Mason, *acta physica slovacica* **56(4)** (2006) 521.
- [8] G. Staszewska, D. W. Schwenke, D. Thirumalai, D. G. Truhlar, Phys. Rev. A **28** (1983) 2740.
- [9] F. Blanco, G. Garcia, Phys. Rev. A **67**, (2003) 022701.
- [10] G. Garcia, F. Blanco, Phys. Rev. A **62**, (2000) 044702.
- [11] Y.-K. Kim and P.M. Stone, Phys. Rev. A **64**, (2001) 052707.
- [12] D Margreiter, H Deutsch, TD Maerk, Int J Mass Spectrom Ion Processes **139**, (1994) 127.
- [13] W. Lotz, Z. Phys. **232**, 101 (1970), J. Opt. Soc. Am. **60**, (1970) 206.
- [14] R. S. Freund, R C Wetzal, R J Shul, and T R Hayes, Phys Rev. A **41**, (1990) 3575.

# PRODUCTION AND PURIFICATION OF BIOETHANOL BY FERMENTATION PROCESS USING VARIOUS CHEAP RAW MATERIALS

A. Umesh Kumar

**Abstract**— In this present work, bioethanol was produced by using 6 various cheap raw biomass. The biomasses used are potato, sweet potato, fruit pulp, paddy straws, paddy husk, and sugar cane beets. To produced ethanol from biomass two key processes were followed. First the starch or hemicelluloses and cellulose portions of the biomass were broken down into simple sugar through a process called saccharification. Second the sugars are fermented to produce ethanol. The fermented product was purified by primary distillation process at 80<sup>o</sup>c and fraction collected. The amount of ethanol is then determined by specific gravity method. Further study can be done to yield more amount of ethanol by optimizing the activity of enzymes over biomass, incubate temperature incubate time and by secondary distillation after primary distillation.

**Index Terms:** --- Amylase, bioethanol, cellulose, fermentation, saccharification, hemicelluloses

## I. INTRODUCTION

The alcohols mainly produced by the microbes by the fermentation process are ethanol and methane, although other potential fuels can be generated using microorganisms, including hydrogen, ethane, propene and butanol. Ethanol has been produced by humans for thousands years. The very first time, ethanol existed only in alcoholic drinks. The usage of ethanol in different fields was highly extended by establishing some new method in the purification process. (Elsheimer., *et al.*, 1999 ).

Ethanol or ethyl alcohol (C<sub>2</sub>H<sub>5</sub>OH) is a clear colorless liquid with a burning taste and spirituous smell that is miscible with water in all proportion. It is biodegradable with low toxicity and cause little environmental pollution if spilt. Bioethanol is a renewable resources with a clean combustion, a high octane number (103), is chemically stable, no toxic and biodegradable.

A. Author is Asst. Prof. At MEFGI Rajkot

## A. Bioethanol

Bioethanol can be defined as the ethanol produced by fermenting sugars that are extracted from agricultural or biodegradable waste using micro organisms. The biodegradable waste like sawdust, waste straws, sugarcane beets, whisky waste, wood chips, rice husk etc is mainly used. The bioethanol are produced by this process are used as a fuel for the transport of vehicles. Bioethanol fuel can be also produced by the chemical process of reacting ethylene with steam. But the production from ethylene is losing importance because of its high cost. The basic concern over bioethanol production expansion is depletion of natural resources and demands for environmentally acceptable fuels – biofuels from renewable feedstock's, emitting less carbon dioxide into the atmosphere.

## B. Advantages of bioethanol

The advantage of bioethanol for the environment is its potential to be a carbon neutral on a life cycle basis. The carbon dioxide (CO<sub>2</sub>) emitted during its use is offset by the absorption from the atmosphere during its growth.

The fermentation process proper is proceeded by the production of a suitable fermentation substrate. There are three main substrate used for the bioethanol production ( Brown, William Henry et al., 2000).

- First one is the sugar containing raw material in liquid form.
- Second one is the starch containing raw materials.
- Third one is the cellulose containing raw materials.

The enzymes mainly used for the hydrolysis of substrate are amylase and cellulase because the raw material contain large amount of starch and cellulose. The substrate containing these polysaccharides is first subjected to heat or enzyme treatment to convert these polysaccharides into fermentable sugar. These fermentable sugars are then converted to ethanol by action of yeast.

## II. AIM & OBJECTIVES OF THIS STUDY

- To produce the amylase from potato extract by *Aspergillus* species.
- To produce the cellulase by *Aspergillus* species.
- To produce the bioethanol by fermentation using cheap raw materials.
- Purification of bioethanol.

And to achieve this aim the present strategy will be as below. Presently various cheap raw materials are used for the production of Bioethanol. The raw materials which mainly contain starch or cellulose are utilized by the microorganism for the production of ethanol. A typical biomass contain about 45% of hexose, 30% pentose and 25% other substance. Pentose (mostly xylose) cannot be fermented by traditional ethanol producing yeast. Thus genetically engineered spices are used for the utilization of xylose. However, to make the ethanol production process economically feasible, several research challenges remain in order to further improve the overall yield of both ethanol and solid fuel, increase the productivity in the conversion steps and to reduce the production cost.

The main challenges are:

- To improve enzymatic hydrolysis and especially the production of cellulases
- To improve fermentation of all sugars available in wood and to make the fermenting organism more tolerant to inhibitors
- To increase process integration to reduce the number of process steps, the energy demand and to re-use process streams in order to minimize the use of fresh water and reduce the amount of waste streams.

## III. MATERIALS AND METHOD

There are 6 types of samples were collected these are the followings;

- Potato
- Sweet potato
- Fruit pulp.
- Sugar cane beets
- Paddy straws
- Paddy husk

Samples were collected from local market of Domlur and Karnataka.

### A. Production of enzymes

- Two enzymes were used for hydrolysis the samples. 3 samples (Potato, sweet potato, and fruit pulp ) were hydrolyzed by amylase and the other 3 samples (sugar

cane beets , Paddy straws and Paddy husk) were hydrolyzed by cellulase

### A. Production of Amylase:

50g of potato were taken and peel the skin. Then cut the potato into small pieces. The small pieces of potato were taken in a beaker and boil it with 150 ml of distilled water for 20 to 30 min. After boiling the potato was taken in a pistil and smash it well, filter through a muslin cloth into a clean 250 mL Erlenmeyer flask and it then was then autoclave for 15 min at 15 lbs. Allow it to cool and inoculated a loop full culture of *Aspergillus awomuri* and kept in the shaker for about 72 hrs at room temperature. After the incubation the culture was centrifuged at 5000 rpm for 15 minutes. The supernatant was taken and transferred into a clean conical flask. Extraction was done under sterile conditions to prevent any microbial contamination.. The clear supernatant broth was used to determine the enzyme yield.

### B. Production of Cellulase:

50g Paddy straw was taken in a conical flask and added 150ml distilled water and boiled for 30min. After boiling the extract was filtered. From this extract 150 ml was taken and then prepared production media, then kept for autoclave for 50min and allowed to cool. After cooling a loop full of culture (*Aspergillus awomuri*) was inoculated into the medium in 250 mL Erlenmeyer flask and incubated at room temperature for about 72 hours. After incubation the flask were retrieved from the shaker. Cultural broth was transferred to sterile centrifuge tubes, centrifuged at 5000 rpm for 15 minutes. The supernatant was taken and transferred into a clean conical flask. Extraction was done under sterile conditions to prevent any microbial contamination. The crude filtrate containing enzyme is assayed for its activity in terms of filter paper units through DNS method.

Enzymatically Hydrolysed Substrate was followed by fermentation & distillation for further utilization.

## IV. OBSERVATIONS & RESULTS

### A. Assay of amylase

In 72 hrs of incubation at 37o C and pH -7, when the supernatant was taken for assay, instance blue color was observed in the test tube marked 0, 0.2, 0.4 and the intensity of the color was decreasing in the test tube marked as 0.6 and 0.8. Only light blue color was observed in the test tube marked as 1.0 and no color was observed in the test tube marked as 2.0

Graphical representation of the assay

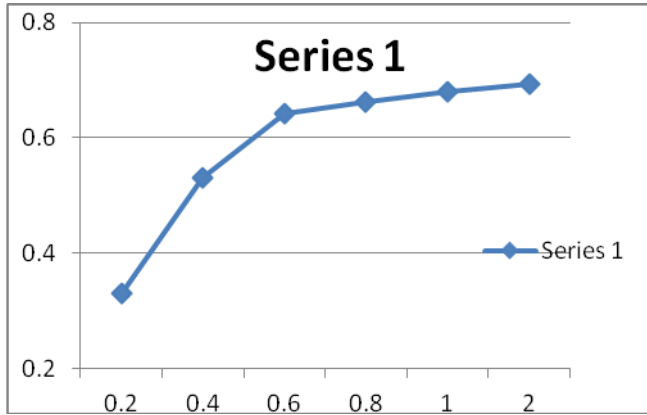
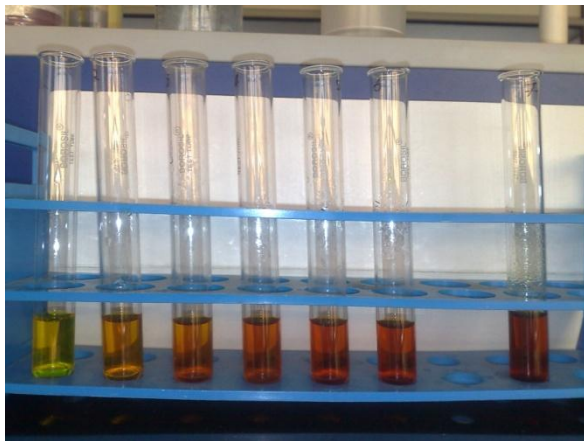
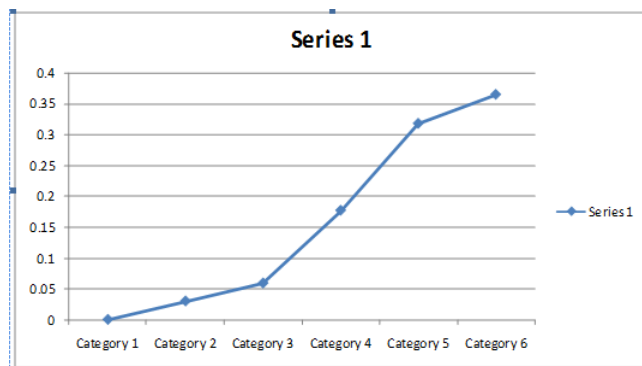


Figure:-Shows amylase assays in different concentrations

B. Assay for cellulose

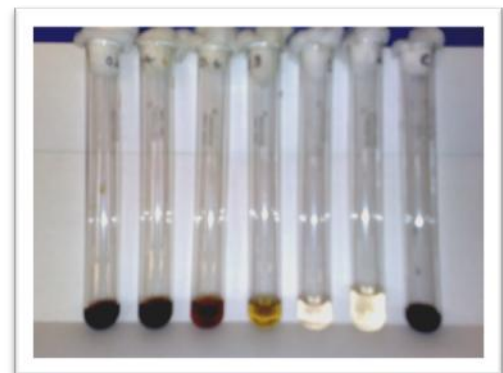


Graphical representation of Assay



Results

SL.N	Sample	Volume of extract before distillation in ml	Volume of extract after distillation in ml	Volume of Bioethanol after distillation in ml
1	Paddy husk	150	148	1.5
2	Paddy straws	150	148	1.5
3	Sweet potato	150	149	1.5
4	Potato	150	148	1.0
5	Fruit extract	150	146	1.5
6	Sugar cane beets	150	146	3.5





**V. REFERENCE**

- [1] Akpan, I., Uraih, N., Obuekwe, C. O. and Ikenebomeh, M. J. (2004) "Production of ethanol From cassava waste" *Acta Biotechnological* 8(1) 39-
- [2] AZZAM, A.M., 1989, Pretreatment of can bagasse with alkaline hydrogen peroxide for Enzymatic hydrolysis of cellulose and ethanol fermentation. *Journal of Environmental Science*, 24: 421-4
- [3] Arthe, R., R, Rajesh., E, M.Rajesh., R, Rajendran., S, Jeyachandran., (2008).
- [4] Arja Miettinen Oinonen, Banat IM, Nigam P, Singh D, Marchant P, McHale AP (1998) Ethanol production at elevated temperatures and alcohol concentrations. Part I: Yeasts in general. *World J Microbial Biotechnology* 14:809–821
-

# Research Methodology

A. Nirupama Patel

## Abstract

Research in common parlance refers to a search for knowledge. One can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation. The Advanced Learner's Dictionary of Current English lays down the meaning of research as "a careful investigation or inquiry specially through search for new facts in any branch of knowledge." [1]. Redman and Mory define research as a "systematized effort to gain new knowledge." [2]. Some people consider research as a movement, a movement from the known to the unknown. It is actually a voyage of discovery. We all possess the vital instinct of inquisitiveness for, when the unknown confronts us, we wonder and our inquisitiveness makes us probe and attain full and fuller understanding of the unknown. This inquisitiveness is the mother of all knowledge and the method, which man employs for obtaining the knowledge of whatever the unknown, can be termed as research.

## I INTRODUCTION

Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis. D. Slesinger and M. Stephenson in the Encyclopedia of Social Sciences define research as "the manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art." [3]. Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the pursuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research. The systematic approach concerning generalization and the formulation of a theory is also research. As such the term 'research' refers to the

systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analyzing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalizations for some theoretical formulation.

## II OBJECTIVES OF RESEARCH

The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet. Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:

1. To gain familiarity with a phenomenon or to achieve new insights into it.
2. To portray accurately the characteristics of a particular individual, situation or a group.
3. To determine the frequency with which something occurs or with which it is associated with something else.
4. To test a hypothesis of a causal relationship between viable.

## III MOTIVATION IN RESEARCH

What makes people to undertake research? This is a question of fundamental importance. The possible motives for doing research may be either one or more of the following:

1. Desire to get a research degree along with its consequential benefits.
2. Desire to face the challenge in solving the unsolved problems, i.e., concern over practical problems initiates research.
3. Desire to get intellectual joy of doing some creative work.
4. Desire to be of service to society.
5. Desire to get respectability.

However, this is not an exhaustive list of factors motivating people to undertake research studies. Many more factors such as directives of government, employment conditions, curiosity about new things, desire to understand causal relationships, social thinking

and awakening, and the like may as well motivate people to perform research operations.

#### IV TYPES OF RESEARCH

##### 1. *Descriptive vs. Analytical:*

*Descriptive research* includes surveys and fact-finding enquiries of different kinds. The major purpose of descriptive research is description of the state of affairs as it exists at present. In social science and business research we quite often use the term *Ex post facto research* for descriptive research studies. The main characteristic of this method is that the researcher has no control over the variables; he can only report what has happened or what is happening. Most *ex post facto research* projects are used for descriptive studies in which the researcher seeks to measure such items as, for example, frequency of shopping, preferences of people, or similar data. *Ex post facto studies* also include attempts by researchers to discover causes even when they cannot control the variables. The methods of research utilized in descriptive research are survey methods of all kinds, including comparative and correlation methods. In *analytical research*, on the other hand, the researcher has to use facts or information already available, and analyze these to make a critical evaluation of the material.

##### 2. *Applied vs. Fundamental:*

Research can either be applied (or action) research or fundamental (to basic or pure) research. *Applied research* aims at finding a solution for an immediate problem facing a society or an industrial/business organization, whereas *fundamental research* is mainly concerned with generalizations and with the formulation of a theory. "Gathering knowledge for knowledge's sake is termed 'pure' or 'basic' research." [4]. Research concerning some natural phenomenon or relating to pure mathematics are examples of fundamental research. Similarly, research studies, concerning human behaviour carried on with a view to make generalizations about human behaviour, are also examples of fundamental research, but research aimed at certain conclusions (say, a solution) facing a concrete social or business problem is an example of applied research. Research to identify social, economic or political trends that may affect a particular institution or the copy research (research to find out whether certain communications will be read and understood) or the marketing research or evaluation research are examples of applied research. Thus, the central aim of applied research is to discover a solution for some pressing practical problem, whereas basic research is directed towards finding information that has a broad base of applications and thus, adds to the already existing organized body of scientific knowledge.

##### 3. *Quantitative vs. Qualitative:*

Quantitative research is based on the measurement of quantity or amount. It is applicable to phenomena that can be expressed in terms of quantity. Qualitative research, on the other hand, is concerned with qualitative phenomenon, i.e., phenomena relating to or involving quality or kind. For instance, when we are interested in investigating the reasons for human behaviour (i.e., why people think or do certain things), we quite often talk of 'Motivation Research', an important type of qualitative research. This type of research aims at discovering the underlying motives and desires, using in depth interviews for the purpose. Other techniques of such research are word association tests, sentence completion tests, story completion tests and similar other projective techniques. Attitude or opinion research i.e., research designed to find out how people feel or what they think about a particular subject or institution is also qualitative research. Qualitative research is specially important in the behavioural sciences where the aim is to discover the underlying motives of human behaviour. Through such research we can analyse the various factors which motivate people to behave in a particular manner or which make people like or dislike a particular thing. It may be stated, however, that to apply qualitative research in practice is relatively a difficult job and therefore, while doing such research, one should seek guidance from experimental psychologists.

##### 4. *Conceptual vs. Empirical:*

Conceptual research is that related to some abstract idea(s) or theory. It is generally used by philosophers and thinkers to develop new concepts or to reinterpret existing ones. On the other hand, empirical research relies on experience or observation alone, often without due regard for system and theory. It is data-based research, coming up with conclusions which are capable of being verified by observation or experiment. We can also call it as experimental type of research. In such a research it is necessary to get at facts firsthand, at their source, and actively to go about doing certain things to stimulate the production of desired information. In such a research, the researcher must first provide himself with a working hypothesis or guess as to the probable results. He then works to get enough facts (data) to prove or disprove his hypothesis. He then sets up experimental designs which he thinks will manipulate the persons or the materials concerned so as to bring forth the desired information. Such research is thus characterized by the experimenter's control over the variables under study and his deliberate manipulation of one of them to study its effects. Empirical research is appropriate when proof is sought that certain variables affect other variables in some way. Evidence gathered through experiments or empirical studies is today considered to be the most powerful support possible for a given hypothesis.

### 5. Some Other Types of Research:

All other types of research are variations of one or more of the above stated approaches, based on either the purpose of research, or the time required to accomplish research, on the environment in which research is done, or on the basis of some other similar factor. From the point of view of time, we can think of research either as *one-time research* or *longitudinal research*. In the former case the research is confined to a single time-period, whereas in the latter case the research is carried on over several time-periods. Research can be *field-setting research* or *laboratory research* or *simulation research*, depending upon the environment in which it is to be carried out. Research can as well be understood as *clinical* or *diagnostic research*. Such research follow case-study methods or indepth approaches to reach the basic causal relations. Such studies usually go deep into the causes of things or events that interest us, using very small samples and very deep probing data gathering devices. The research may be *exploratory* or it may be formalized. The objective of exploratory research is the development of hypotheses rather than their testing, whereas formalized research studies are those with substantial structure and with specific hypotheses to be tested. *Historical research* is that which utilizes historical sources like documents, remains, etc. to study events or ideas of the past, including the philosophy of persons and groups at any remote point of time. Research can also be classified as *conclusion-oriented* and *decision-oriented*. While doing conclusion- oriented research, a researcher is free to pick up a problem, redesign the enquiry as he proceeds and is prepared to conceptualize as he wishes. Decision-oriented research is always for the need of a decision maker and the researcher in this case is not free to embark upon research according to his own inclination. Operations research is an example of decision oriented research since it is a scientific method of providing executive departments with a quantitative basis for decisions regarding operations under their control.

### V RESEARCH APPROCHES

The above description of the types of research brings to light the fact that there are two basic approaches to research, viz., *quantitative approach* and the *qualitative approach*. The former involves the generation of data in quantitative form which can be subjected to rigorous quantitative analysis in a formal and rigid fashion. This approach can be further sub-classified into *inferential*, *experimental* and *simulation approaches* to research. The purpose of *inferential approach* to research is to form a data base from which to infer characteristics or relationships of population. This usually means survey research where a sample of population is studied (questioned or observed) to determine its characteristics, and it is then inferred that the population has the same characteristics. *Experimental*

*approach* is characterized by much greater control over the research environment and in this case some variables are manipulated to observe their effect on other variables. *Simulation approach* involves the construction of an artificial environment within which relevant information and data can be generated. This permits an observation of the dynamic behaviour of a system (or its sub-system) under controlled conditions. The term ‘simulation’ in the context of business and social sciences applications refers to “the operation of a numerical model that represents the structure of a dynamic process. Given the values of initial conditions, parameters and exogenous variables, a simulation is run to represent the behaviour of the process over time.” [5]. Simulation approach can also be useful in building models for understanding future conditions.

*Qualitative approach* to research is concerned with subjective assessment of attitudes, opinions and behaviour. Research in such a situation is a function of researcher’s insights and impressions. Such an approach to research generates results either in non-quantitative form or in the form which are not subjected to rigorous quantitative analysis. Generally, the techniques of focus group interviews, projective techniques and depth interviews are used. All these are explained at length in chapters that follow.

### VI SIGNIFICANCE OF RESEARCH

All progress is born of inquiry. Doubt is often better than overconfidence, for it leads to inquiry, and inquiry leads to invention” which the significance of research can well be understood. Increased amounts of research make progress possible. Research inculcates scientific and inductive thinking and it promotes the development of logical habits of thinking and organization.

The role of research in several fields of applied economics, whether related to business or to the economy as a whole, has greatly increased in modern times. The increasingly complex nature of business and government has focused attention on the use of research in solving operational problems. Research, as an aid to economic policy, has gained added importance, both for government and business.

Research provides the basis for nearly all government policies in our economic system. For instance, government’s budgets rest in part on an analysis of the needs and desires of the people and on the availability of revenues to meet these needs. The cost of needs has to be equated to probable revenues and this is a field where research is most needed. Through research we can devise alternative policies and can as well examine the consequences of each of these alternatives.

Decision-making may not be a part of research, but research certainly facilitates the decisions of the policy maker. Government has also to chalk out programmes for dealing with all facets of the country’s existence and most of

these will be related directly or indirectly to economic conditions. The plight of cultivators, the problems of big and small business and industry, working conditions, trade union activities, the problems of distribution, even the size and nature of defence services are matters requiring research. Thus, research is considered necessary with regard to the allocation of nation's resources. Another area in government, where research is necessary, is collecting information on the economic and social structure of the nation. Such information indicates what is happening in the economy and what changes are taking place. Collecting such statistical information is by no means a routine task, but it involves a variety of research problems. These days nearly all governments maintain large staff of research technicians or experts to carry on this work. Thus, in the context of government, research as a tool to economic policy has three distinct phases of operation, viz., (i) investigation of economic structure through continual compilation of facts; (ii) diagnosis of events that are taking place and the analysis of the forces underlying them; and (iii) the prognosis, i.e., the prediction of future developments.

Research has its special significance in solving various operational and planning problems of business and industry. Operations research and market research, along with motivational research, are considered crucial and their results assist, in more than one way, in taking business decisions. Market research is the investigation of the structure and development of a market for the purpose of formulating efficient policies for purchasing, production and sales. Operations research refers to the application of mathematical, logical and analytical techniques to the solution of business problems of cost minimization or of profit maximization or what can be termed as optimization problems. Motivational research of determining why people behave as they do is mainly concerned with market characteristics. In other words, it is concerned with the determination of motivations underlying the consumer (market) behaviour. All these are of great help to people in business and industry who are responsible for taking business decisions. Research with regard to demand and market factors has great utility in business. Given knowledge of future demand, it is generally not difficult for a firm, or for an industry to adjust its supply schedule within the limits of its projected capacity. Market analysis has become an integral tool of business policy these days. Business budgeting, which ultimately results in a projected profit and loss account, is based mainly on sales estimates which in turn depends on business research. Once sales forecasting is done, efficient production and investment programmes can be set up around which are grouped the purchasing and financing plans. Research, thus, replaces intuitive business decisions by more logical and scientific decisions.

Research is equally important for social scientists in studying social relationships and in seeking answers to various social problems. It provides the intellectual satisfaction of knowing a few things just for the sake of

knowledge and also has practical utility for the social scientist to know for the sake of being able to do something better or in a more efficient manner. Research in social sciences is concerned both with knowledge for its own sake and with knowledge for what it can contribute to practical concerns. "This double emphasis is perhaps especially appropriate in the case of social science. On the one hand, its responsibility as a science is to develop a body of principles that make possible the understanding and prediction of the whole range of human interactions. On the other hand, because of its social orientation, it is increasingly being looked to for practical guidance in solving immediate problems of human relations." [6].

In addition to what has been stated above, the significance of research can also be understood keeping in view the following points:

- (a) To those students who are to write a master's or Ph.D. thesis, research may mean a careerism or a way to attain a high position in the social structure;
- (b) To professionals in research methodology, research may mean a source of livelihood;
- (c) To philosophers and thinkers, research may mean the outlet for new ideas and insights;
- (d) To literary men and women, research may mean the development of new styles and creative work;
- (e) To analysts and intellectuals, research may mean the generalizations of new theories.

Thus, research is the fountain of knowledge for the sake of knowledge and an important source of providing guidelines for solving different business, governmental and social problems. It is a sort of formal training which enables one to understand the new developments in one's field in a better way.

## VII RESEARCH METHODS VERSUS METHODOLOGY

It seems appropriate at this juncture to explain the difference between research methods and research methodology. *Research methods* may be understood as all those methods/techniques that are used for conduction of research. Research methods or techniques, thus, refer to the methods the researchers use in performing research operations. In other words, all those methods which are used by the researcher during the course of studying his research problem are termed as research methods. Since the object of research, particularly the applied research, it to arrive at a solution for a given problem, the available data and the unknown aspects of the problem have to be related to each other to make a solution possible. Keeping this in view, research methods can be put into the following three groups:

1. In the first group we include those methods which are concerned with the collection of data. These methods will be used where the data already available are not sufficient to arrive at the required solution;
2. The second group consists of those statistical techniques which are used for establishing relationships between the data and the unknowns;
3. The third group consists of those methods which are used to evaluate the accuracy of the results obtained.

Research methods falling in the above stated last two groups are generally taken as the analytical tools of research.

Research methodology is a way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically. In it we study the various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them. It is necessary for the researcher to know not only the research methods/techniques but also the methodology. Researchers not only need to know how to develop certain indices or tests, how to calculate the mean, the mode, the median or the standard deviation or chi-square, how to apply particular research techniques, but they also need to know which of these methods or techniques, are relevant and which are not, and what would they mean and indicate and why. Researchers also need to understand the assumptions underlying various techniques and they need to know the criteria by which they can decide that certain techniques and procedures will be applicable to certain problems and others will not. All this means that it is necessary for the researcher to design his methodology for his problem as the same may differ from problem to problem. For example, an architect, who designs a building, has to consciously evaluate the basis of his decisions, i.e., he has to evaluate why and on what basis he selects particular size, number and location of doors, windows and ventilators, uses particular materials and not others and the like. Similarly, in research the scientist has to expose the research decisions to evaluation before they are implemented. He has to specify very clearly and precisely what decisions he selects and why he selects them so that they can be evaluated by others also.

From what has been stated above, we can say that research methodology has many dimensions and research methods do constitute a part of the research methodology. The scope of research methodology is wider than that of research methods. Thus, when we talk of research methodology we not only talk of the research methods but also consider the logic behind the methods we use in the context of our research study and explain why we are using a particular method or technique and why we are not using others so that research results are capable of being evaluated either by the researcher himself or by others. Why a

research study has been undertaken, how the research problem has been defined, in what way and why the hypothesis has been formulated, what data have been collected and what particular method has been adopted, why particular technique of analyzing data has been used and a host of similar other questions are usually answered when we talk of research methodology concerning a research problem or study.

#### VIII RESEARCH PROCESS

Before embarking on the details of research methodology and techniques, it seems appropriate to present a brief overview of the research process. Research process consists of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps.

One should remember that the various steps involved in a research process are not mutually exclusive; nor they are separate and distinct. They do not necessarily follow each other in any specific order and the researcher has to be constantly anticipating at each step in the research process the requirements of the subsequent steps. However, the following order concerning various steps provides a useful procedural guideline regarding the research process: (1) formulating the research problem; (2) extensive literature survey; (3) developing the hypothesis; (4) preparing the research design; (5) determining sample design; (6) collecting the data; (7) execution of the project; (8) analysis of data; (9) hypothesis testing; (10) generalizations and interpretation, and (11) preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached.

A brief description of the above stated steps will be helpful.

##### *1. Formulating the research problem:*

There are two types of research problems, viz., those which relate to states of nature and those which relate to relationships between variables. At the very outset the researcher must single out the problem he wants to study, i.e., he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into. Initially the problem may be stated in a broad general way and then the ambiguities, if any, relating to the problem be resolved. Then, the feasibility of a particular solution has to be considered before a working formulation of the problem can be set up. The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific enquiry. Essentially two steps are involved in formulating the research problem, viz., understanding the problem thoroughly, and rephrasing the same into meaningful terms from an analytical point of view.

The best way of understanding the problem is to discuss it with one's own colleagues or with those

having some expertise in the matter. In an academic institution the researcher can seek the help from a guide who is usually an experienced man and has several research problems in mind. Often, the guide puts forth the problem in general terms and it is up to the researcher to narrow it down and phrase the problem in operational terms. In private business units or in governmental organizations, the problem is usually earmarked by the administrative agencies with whom the researcher can discuss as to how the problem originally came about and what considerations are involved in its possible solutions.

The researcher must at the same time examine all available literature to get himself acquainted with the selected problem. He may review two types of literature—the conceptual literature concerning the concepts and theories, and the empirical literature consisting of studies made earlier which are similar to the one proposed. The basic outcome of this review will be the knowledge as to what data and other materials are available for operational purposes which will enable the researcher to specify his own research problem in a meaningful context. After this the researcher rephrases the problem into analytical or operational terms i.e., to put the problem in as specific terms as possible. This task of formulating, or defining, a research problem is a step of greatest importance in the entire research process. The problem to be investigated must be defined unambiguously for that will help discriminating relevant data from irrelevant ones. Care must, however, be taken to verify the objectivity and validity of the background facts concerning the problem. The statement of the objective is of basic importance because it determines the data which are to be collected, the characteristics of the data which are relevant, relations which are to be explored, the choice of techniques to be used in these explorations and the form of the final report. If there are certain pertinent terms, the same should be clearly defined along with the task of formulating the problem. In fact, formulation of the problem often follows a sequential pattern where a number of formulations are set up, each formulation more specific than the preceding one, each one phrased in more analytical terms, and each more realistic in terms of the available data and resources.

### *2. Extensive literature survey:*

Once the problem is formulated, a brief summary of it should be written down. It is compulsory for a research worker writing a thesis for a Ph.D. degree to write a synopsis of the topic and submit it to the necessary Committee or the Research Board for approval. At this juncture the researcher should undertake extensive literature survey connected with the problem. For this purpose, the abstracting and indexing journals and published or unpublished bibliographies are the first place to go to. Academic journals, conference

proceedings, government reports, books etc., must be tapped depending on the nature of the problem. In this process, it should be remembered that one source will lead to another. The earlier studies, if any, which are similar to the study in hand should be carefully studied. A good library will be a great help to the researcher at this stage.

### *3. Development of working hypotheses:*

After extensive literature survey, researcher should state in clear terms the working hypothesis or hypotheses. Working hypothesis is tentative assumption made in order to draw out and test its logical or empirical consequences. As such the manner in which research hypotheses are developed is particularly important since they provide the focal point for research. They also affect the manner in which tests must be conducted in the analysis of data and indirectly the quality of data which is required for the analysis. In most types of research, the development of working hypothesis plays an important role. Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested. The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track. It sharpens his thinking and focuses attention on the more important facets of the problem. It also indicates the type of data required and the type of methods of data analysis to be used.

How does one go about developing working hypotheses? The answer is by using the following approach:

- (a) Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;
- (b) Examination of data and records, if available, concerning the problem for possible trends, peculiarities and other clues;
- (c) Review of similar studies in the area or of the studies on similar problems; and
- (d) Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.

Thus, working hypotheses arise as a result of a-priori thinking about the subject, examination of the available data and material including related studies and the counsel of experts and interested parties. Working hypotheses are more useful when stated in precise and clearly defined terms. It may as well be remembered that occasionally we may encounter a problem where we do not need working hypotheses, specially in the case of exploratory or formulative researches which do not aim at testing the hypothesis. But as a general rule, specification of working hypotheses in another basic step of the research process in most research problems.

### *4. Preparing the research design:*

The research problem having been formulated in clear cut terms, the researcher will be required to prepare a research design, i.e., he will have to state the conceptual structure within which research would be conducted. The preparation of such a design facilitates research to be as efficient as possible yielding maximal information. In other words, the function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money. But how all these can be achieved depends mainly on the research purpose. Research purposes may be grouped into four categories, viz., (i) Exploration, (ii) Description, (iii) Diagnosis, and (iv) Experimentation. A flexible research design which provides opportunity for considering many different aspects of a problem is considered appropriate if the purpose of the research study is that of exploration. But when the purpose happens to be an accurate description of a situation or of an association between variables, the suitable design will be one that minimizes bias and maximizes the reliability of the data collected and analysed.

There are several research designs, such as, experimental and non-experimental hypothesis testing. Experimental designs can be either informal designs (such as before-and-after without control, after-only with control, before-and-after with control) or formal designs (such as completely randomized design, randomized block design, Latin square design, simple and complex factorial designs), out of which the researcher must select one for his own project.

The preparation of the research design, appropriate for a particular research problem, involves usually the consideration of the following:

- (i) the means of obtaining the information;
- (ii) the availability and skills of the researcher and his staff (if any);
- (iii) explanation of the way in which selected means of obtaining information will be organized and the reasoning leading to the selection;
- (iv) the time available for research; and
- (v) the cost factor relating to research, i.e., the finance available for the purpose.

##### 5. *Determining sample design:*

All the items under consideration in any field of inquiry constitute a 'universe' or 'population'. A complete enumeration of all the items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry when all the items are covered no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases. Moreover, there is no way of checking the element of bias or its extent except

through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Not only this, census inquiry is not possible in practice under many circumstances. For instance, blood testing is done only on sample basis. Hence, quite often we select only a few items from the universe for our study purposes. The items so selected constitute what is technically called a sample.

The researcher must decide the way of selecting a sample or what is popularly known as the sample design. In other words, a sample design is a definite plan determined before any data are actually collected for obtaining a sample from a given population. Thus, the plan to select 12 of a city's 200 drugstores in a certain way constitutes a sample design. Samples can be either probability samples or non-probability samples. With probability samples each element has a known probability of being included in the sample but the non-probability samples do not allow the researcher to determine this probability. Probability samples are those based on simple random sampling, systematic sampling, stratified sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgment sampling and quota sampling techniques. A brief mention of the important sample designs is as follows:

##### a). *Deliberate sampling:*

Deliberate sampling is also known as purposive or non-probability sampling. This sampling method involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe. When population elements are selected for inclusion in the sample based on the ease of access, it can be called *convenience sampling*. If a researcher wishes to secure data from, say, gasoline buyers, he may select a fixed number of petrol stations and may conduct interviews at these stations. This would be an example of convenience sample of gasoline buyers. At times such a procedure may give very biased results particularly when the population is not homogeneous. On the other hand, in *judgement sampling* the researcher's judgement is used for selecting items which he considers as representative of the population. For example, a judgement sample of college students might be taken to secure reactions to a new method of teaching. Judgement sampling is used quite frequently in qualitative research where the desire happens to be to develop hypotheses rather than to generalize to larger populations.

##### b) *Simple random sampling:*

This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected. For example, if we have to select a sample of 300 items from a universe of 15,000 items, then we can put the names or



numbers of all the 15,000 items on slips of paper and conduct a lottery. Using the random number tables is another method of random sampling. To select the sample, each item is assigned a number from 1 to 15,000. Then, 300 five digit random numbers are selected from the table. To do this we select some random starting point and then a systematic pattern is used in proceeding through the table. We might start in the 4th row, second column and proceed down the column to the bottom of the table and then move to the top of the next column to the right. When a number exceeds the limit of the numbers in the frame, in our case over 15,000, it is simply passed over and the next number selected that does fall within the relevant range. Since the numbers were placed in the table in a completely random fashion, the resulting sample is random. This procedure gives each item an equal probability of being selected. In case of infinite population, the selection of each item in a random sample is controlled by the same probability and that successive selections are independent of one another.

c) *Systematic sampling:*

In some instances the most practical way of sampling is to select every 15th name on a list, every 10th house on one side of a street and so on. Sampling of this type is known as systematic sampling. An element of randomness is usually introduced into this kind of sampling by using random numbers to pick up the unit with which to start. This procedure is useful when sampling frame is available in the form of a list. In such a design the selection process starts by picking some random point in the list and then every  $n$ th element is selected until the desired number is secured.

D) *Stratified sampling:*

If the population from which a sample is to be drawn does not constitute a homogeneous group, then stratified sampling technique is applied so as to obtain a representative sample. In this technique, the population is stratified into a number of non-overlapping subpopulations or strata and sample items are selected from each stratum. If the items selected from each stratum is based on simple random sampling the entire procedure, first stratification and then simple random sampling, is known as *stratified random sampling*.

e) *Quota sampling:*

In stratified sampling the cost of taking random samples from individual strata is often so expensive that interviewers are simply given quota to be filled from different strata, the actual selection of items for sample being left to the interviewer's judgement. This is called quota sampling. The size of the quota for each stratum is generally proportionate to the size of that stratum in the population. Quota sampling is thus an important form of non-

probability sampling. Quota samples generally happen to be judgement samples rather than random samples.

f) *Cluster sampling and area sampling:*

Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample. Suppose some departmental store wishes to sample its credit card holders. It has issued its cards to 15,000 customers. The sample size is to be kept say 450. For cluster sampling this list of 15,000 card holders could be formed into 100 clusters of 150 card holders each. Three clusters might then be selected for the sample randomly. The sample size must often be larger than the simple random sample to ensure the same level of accuracy because in cluster sampling procedural potential for order bias and other sources of error is usually accentuated. The clustering approach can, however, make the sampling procedure relatively easier and increase the efficiency of field work, especially in the case of personal interviews.

*Area sampling* is quite close to cluster sampling and is often talked about when the total geographical area of interest happens to be big one. Under area sampling we first divide the total area into a number of smaller non-overlapping areas, generally called geographical clusters, then a number of these smaller areas are randomly selected, and all units in these small areas are included in the sample. Area sampling is specially helpful where we do not have the list of the population concerned. It also makes the field interviewing more efficient since interviewer can do many interviews at each location.

G) *Multi-stage sampling:*

This is a further development of the idea of cluster sampling. This technique is meant for big inquiries extending to a considerably large geographical area like an entire country. Under multi-stage sampling the first stage may be to select large primary sampling units such as states, then districts, then towns and finally certain families within towns. If the technique of random-sampling is applied at all stages, the sampling procedure is described as multi-stage random sampling.

h) *Sequential sampling:*

This is somewhat a complex sample design where the ultimate size of the sample is not fixed in advance but is determined according to mathematical decisions on the basis of information yielded as survey progresses. This design is usually adopted under acceptance sampling plan in the context of statistical quality control.

In practice, several of the methods of sampling described above may well be used in the same study in which case it can be called mixed sampling. It may be pointed out here that normally one should resort to random sampling so that bias can be eliminated and sampling error can be estimated. But purposive sampling

is considered desirable when the universe happens to be small and a known characteristic of it is to be studied intensively. Also, there are conditions under which sample designs other than random sampling may be considered better for reasons like convenience and low costs. The sample design to be used must be decided by the researcher taking into consideration the nature of the inquiry and other related factors.

#### *6. Collecting the data:*

In dealing with any real life problem it is often found that data at hand are inadequate, and hence, it becomes necessary to collect data that are appropriate. There are several ways of collecting the appropriate data which differ considerably in context of money costs, time and other resources at the disposal of the researcher.

Primary data can be collected either through experiment or through survey. If the researcher conducts an experiment, he observes some quantitative measurements, or the data, with the help of which he examines the truth contained in his hypothesis. But in the case of a survey, data can be collected by any one or more of the following ways:

##### *(i) By observation:*

This method implies the collection of information by way of investigator's own observation, without interviewing the respondents. The information obtained relates to what is currently happening and is not complicated by either the past behaviour or future intentions or attitudes of respondents. This method is no doubt an expensive method and the information provided by this method is also very limited. As such this method is not suitable in inquiries where large samples are concerned.

##### *(ii) Through personal interview:*

The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews. This method of collecting data is usually carried out in a structured way where output depends upon the ability of the interviewer to a large extent.

##### *(iii) Through telephone interviews:*

This method of collecting information involves contacting the respondents on telephone itself. This is not a very widely used method but it plays an important role in industrial surveys in developed regions, particularly, when the survey has to be accomplished in a very limited time.

##### *(iv) By mailing of questionnaires:*

The researcher and the respondents do come in contact with each other if this method of survey is adopted. Questionnaires are mailed to the respondents with a request to return after completing the same. It is the most extensively used method in various economic and business surveys. Before applying this method, usually a Pilot Study for testing the questionnaire is conducted which reveals

the weaknesses, if any, of the questionnaire. Questionnaire to be used must be prepared very carefully so that it may prove to be effective in collecting the relevant information.

##### *(v) Through schedules:*

Under this method the enumerators are appointed and given training. They are provided with schedules containing relevant questions. These enumerators go to respondents with these schedules. Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents. Much depends upon the capability of enumerators so far as this method is concerned. Some occasional field checks on the work of the enumerators may ensure sincere work.

The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, financial resources, available time and the desired degree of accuracy. Though he should pay attention to all these factors but much depends upon the ability and experience of the researcher. In collection of statistical data commonsense is the chief requisite and experience the chief teacher.

#### *7. Execution of the project:*

Execution of the project is a very important step in the research process. If the execution of the project proceeds on correct lines, the data to be collected would be adequate and dependable. The researcher should see that the project is executed in a systematic manner and in time. If the survey is to be conducted by means of structured questionnaires, data can be readily machine-processed. In such a situation, questions as well as the possible answers may be coded. If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers. The training may be given with the help of instruction manuals which explain clearly the job of the interviewers at each step. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors in order to keep the survey as much realistic as possible. This, in other words, means that steps should be taken to ensure that the survey is under statistical control so that the collected information is in accordance with the pre-defined standard of accuracy. If some of the respondents do not cooperate, some suitable methods should be designed to tackle this problem. One method of dealing with the non-response problem is to make a list of the non-respondents and take a small sub-sample of them, and then with the help of experts vigorous efforts can be made for securing response.

#### *8. Analysis of data:*

After the data have been collected, the researcher turns to the task of analyzing them. The analysis of data requires

a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, researcher should classify the raw data into some purposeful and usable categories. *Coding* operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted. *Editing* is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation. *Tabulation* is a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture. A great deal of data, specially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study large number of variables affecting a problem simultaneously.

Analysis work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well defined statistical formulae. In the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to tests of significance to determine with what validity data can be said to indicate any conclusion(s). For instance, if there are two samples of weekly wages, each sample being drawn from factories in different parts of the same city, giving two different mean values, then our problem may be whether the two mean values are significantly different or the difference is just a matter of chance. Through the use of statistical tests we can establish whether such a difference is a real one or is the result of random fluctuations. If the difference happens to be real, the inference will be that the two samples come from different universes and if the difference is due to chance, the conclusion would be that the two samples belong to the same universe. Similarly, the technique of analysis of variance can help us in analyzing whether three or more varieties of seeds grown on certain fields yield significantly different results or not. In brief, the researcher can analyze the collected data with the help of various statistical measures.

#### 9. Hypothesis-testing:

After analyzing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier. Do the facts support the hypotheses or they happen to be contrary? This is the usual question which should be answered while testing hypotheses. Various tests, such as Chi square test, *t*-test, *F*-test, have been developed by statisticians for the purpose. The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of research inquiry. Hypothesis-testing will result in either

accepting the hypothesis or in rejecting it. If the researcher had no hypotheses to start with, generalizations established on the basis of data may be stated as hypotheses to be tested by subsequent researches in times to come.

#### 10. Generalizations and interpretation:

If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at generalization, i.e., to build a theory. As a matter of fact, the real value of research lies in its ability to arrive at certain generalizations. If the researcher had no hypothesis to start with, he might seek to explain his findings on the basis of some theory. It is known as interpretation. The process of interpretation may quite often trigger off new questions which in turn may lead to further researches.

#### 11. Preparation of the report or the thesis:

Finally, the researcher has to prepare the report of what has been done by him. Writing of report must be done with great care keeping in view the following:

1. The layout of the report should be as follows: (i) the preliminary pages; (ii) the main text, and (iii) the end matter.

*In its preliminary pages* the report should carry title and date followed by acknowledgements and foreword. Then there should be a table of contents followed by a list of tables and list of graphs and charts, if any, given in the report.

*The main text of the report* should have the following parts:

##### (a) Introduction:

It should contain a clear statement of the objective of the research and an explanation of the methodology adopted in accomplishing the research. The scope of the study along with various limitations should as well be stated in this part.

##### (b) Summary of findings:

After introduction there would appear a statement of findings and recommendations in non-technical language. If the findings are extensive, they should be summarized.

##### (c) Main report:

The main body of the report should be presented in logical sequence and broken-down into readily identifiable sections.

##### (d) Conclusion:

Towards the end of the main text, researcher should again put down the results of his research clearly and precisely. In fact, it is the final summing up.

## IX CRITERIA OF GOOD RESEARCH

Whatever may be the types of research works and studies, one thing that is important is that they all meet on the common ground of scientific method employed by them.

One expects scientific research to satisfy the following criteria: [7]

1. The purpose of the research should be clearly defined and common concepts be used.
2. The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.
3. The procedural design of the research should be carefully planned to yield results that are as objective as possible.
4. The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.
5. The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.
6. Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.
7. Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research and is a person of integrity.

In other words, we can state the qualities of a good research [8] as under:

1. *Good research is systematic:*

It means that research is structured with specified steps to be taken in a specified sequence in accordance with the well defined set of rules. Systematic characteristic of the research does not rule out creative thinking but it certainly does reject the use of guessing and intuition in arriving at conclusions.

2. *Good research is logical:*

This implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research. Induction is the process of reasoning from a part to the whole whereas deduction is the process of reasoning from some premise to a conclusion which follows from that very premise. In fact, logical reasoning makes research more meaningful in the context of decision making.

3. *Good research is empirical:*

It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.

4. *Good research is replicable:*

This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

## X PROBLEMS ENCOUNTERED BY RESEARCHERS IN INDIA

Researchers in India, particularly those engaged in empirical research, are facing several problems. Some of the important problems are as follows:

1. *The lack of a scientific training in the methodology of research* is a great impediment for researchers in our country. There is paucity of competent researchers. Many researchers take a leap in the dark without knowing research methods. Most of the work, which goes in the name of research is not methodologically sound. Research to many researchers and even to their guides, is mostly a scissor and paste job without any insight shed on the collated materials. The consequence is obvious, viz., the research results, quite often, do not reflect the reality or realities. Thus, a systematic study of research methodology is an urgent necessity. Before undertaking research projects, researchers should be well equipped with all the methodological aspects. As such, efforts should be made to provide short- duration intensive courses for meeting this requirement.

2. There is *insufficient interaction* between the university research departments on one side and business establishments, government departments and research institutions on the other side. A great deal of primary data of non-confidential nature remain untouched/untreated by the researchers for want of proper contacts. Efforts should be made to develop satisfactory liaison among all concerned for better and realistic researches. There is need for developing some mechanisms of a university— industry interaction programme so that academics can get ideas from practitioners on what needs to be researched and practitioners can apply the research done by the academics.

3. Most of the business units in our country do not have the confidence that the material supplied by them to researchers will not be misused and as such they are often reluctant in supplying the needed information to researchers. The concept of secrecy seems to be sacrosanct to business organizations in the country so much so that it proves an impermeable barrier to researchers. Thus, there is the need for generating the confidence that the information/data obtained from a business unit will not be misused.

4. *Research studies overlapping one another are undertaken quite often for want of adequate information.* This results in duplication and fritters away resources. This problem can be solved by proper compilation and revision, at regular intervals, of a list of subjects on which and the places where the research is going on. Due attention should be given toward identification of research problems in various disciplines of applied science which are of immediate concern to the industries.

5. *There does not exist a code of conduct for researchers* and inter-university and inter- departmental rivalries are also quite common. Hence, there is need for developing a code of conduct for researchers which, if adhered sincerely, can win over this problem.

6. Many researchers in our country also face *the difficulty of adequate and timely secretarial assistance*, including computerial assistance. This causes unnecessary delays in the completion of research studies. All possible efforts be made in this direction so that efficient secretarial assistance is made available to researchers and that too well in time. University Grants Commission must play a dynamic role in solving this difficulty.

7. *Library management and functioning is not satisfactory at many places* and much of the time and energy of researchers are spent in tracing out the books, journals, reports, etc., rather than in tracing out relevant material from them.

8. *There is also the problem that many of our libraries are not able to get copies of old and new Acts/Rules, reports and other government publications in time.* This problem is felt more in libraries which are away in places from Delhi and/or the state capitals. Thus, efforts should be made for the regular and speedy supply of all governmental publications to reach our libraries.

9. *There is also the difficulty of timely availability of published data* from various government and other agencies doing this job in our country. Researcher also faces the problem on account of the fact that the published data vary quite significantly because of differences in coverage by the concerning agencies.

10. There may, at times, take place *the problem of conceptualization* and also problems relating to the process of data collection and related things.

#### REFERENCES

- [1] *The Advanced Learner's Dictionary of Current English*, Oxford, 1952, p. 1069
- [2] L.V. Redman and A.V.H. Mory, *The Romance of Research*, 1923, p.10.
- [3] *The Encyclopaedia of Social Sciences*, Vol. IX, MacMillan, 1930.
- [4] Pauline V. Young, *Scientific Social Surveys and Research*, p. 30.
- [5] Robert C. Meir, William T. Newell and Harold L. Dazier, *Simulation in Business and Economics*, p. 1.
- [6] Marie Jahoda, Morton Deutsch and Stuart W. Cook, *Research Methods in Social Relations*, p. 4.
- [7] James Harold Fox, *Criteria of Good Research*, Phi Delta Kappan, Vol. 39 (March, 1958), pp. 285–86.
- [8] Danny N. Bellenger and Barnett, A. Greenberg, "Marketing Research—A Management Information Approach", p. 107–108.

# Dimensions of Customer Satisfaction in Services: An Empirical Study on GSM Service Providers

A. Dr. Ritesh K. Patel, B. Mr. Nehal A. Shah

**Abstract** - The Global System of Mobile Communications (GSM) is a second-generation digital technology, which was originally developed in Europe and in less than ten years after the commercial launch, it developed into world's leading and fastest growing mobile standard (GSM Assoc., 2006). The GSM Association estimates that the GSM technology is used by more than one in five people of the world's population, representing approximately 77% of the world's cellular market and is estimated to account for 73% of the world's digital market and 72% of the world's wireless market (GSM Assoc., 2006). This growth principally results from the establishment of new networks in developing countries rather than from an increase in mobile access lines in developed countries (Serenko and Turel, 2006). Asian countries are actively involved in the establishment of the mobile services. Hence it is required to study the services satisfaction of customers with their GSM operators.

**Index Terms** – Customer Satisfaction, Dimensions of Customer Satisfaction, GSM Services, Customer Satisfaction for Services.

## I. INTRODUCTION

The GSM association (Global System for Mobile Communications) was instituted in 1987 to promote and expedite the adoption, development and deployment and evolution of the GSM standard for digital wireless communication. The Association was formed as a result of a European Community agreement on the need to adopt common standards

A. works at Institute of Law, Nirma University, Ahmedabad, 382481, Gujarat, INDIA (e-mail: visit\_ritesh@yahoo.com).

B. Works with Gandhinagar Institute of Technology (A Center for Management Studies), Gandhinagar, 382721, Gujarat, INDIA (e-mail: nehalshahmba@gmail.com).

suitable for cross border European mobile communications.

Indian Telecom sector, like any other industrial sector in the country, has gone through many phases of growth and diversification. Starting from telegraphic and telephonic systems in the 19th century, the field of telephonic communication has now expanded to make use of advanced technologies like GSM, CDMA, and WLL to the great 3G Technology in mobile phones. Day by day, both the Public Players and the Private Players are putting in their resources and efforts to improve the telecommunication technology so as to give the maximum to their customers.

According to the Telecom Regulatory Authority of India (TRAI), the number of telephone subscriber base in the country reached 742.12 million as on October 31, 2010, an increase of 2.61 per cent from 723.28 million in September 2010. With this the overall tele-density (telephones per 100 people) has touched 62.51. The wireless subscriber base has increased to 706.69 million at the end of October 2010 from 687.71 million in September 2010, registering a growth of 2.76 per cent.

Meanwhile, Indian Global System of Mobile Communication (GSM) telecom operators added 17.45 million new subscribers in November 2010, taking the all-India GSM cellular subscriber base to 526.18 million, according to the Cellular Operators Association of India (COAI). The GSM subscriber base stood at 508.72 million at the end of October 2010.

Mobile value added services (VAS) include text or SMS, menu-based services, downloading of music or ring tones, mobile TV, videos and sophisticated m-

commerce applications. As per a report, 'India Telecom 2010' released by KPMG in December 2010, currently, the VAS market is worth US\$ 2.45 billion-US\$ 2.67 billion, which is around 10 per cent of the total revenue of the wireless industry. The share of VAS in wireless revenue is likely to increase to 12-13 per cent by 2011, on the back of increased operator focus on VAS due to continuous fall in voice tariffs, increasing penetration of feature rich handsets, availability of vernacular content and increased user adoption of VAS applications.

## II. LITERATURE REVIEW

### A. Customer Satisfaction

Customer satisfaction, as a construct, has been fundamental to marketing for over three decades.

As early as 1960, Keith (1960) defined marketing as "satisfying the needs and desires of the consumer". Hunt (1982) reported that by the 1970s, interest in customer satisfaction had increased to such an extent that over 500 studies were published. This trend continued and by 1992, Peterson and Wilson estimated the amount of academic and trade articles on customer satisfaction to be over 15,000.

Increasing customer satisfaction has been shown to directly affect companies' market share, which leads to improved profits, positive recommendation, lower marketing expenditures (Reichheld, 1996; Heskett et al., 1997), and greatly impact the corporate image and survival (Pizam and Ellis, 1999).

Taylor and Baker (1994) and Rust and Oliver (1994) identified several factors that precede customer satisfaction and suggested that these factors strongly influence the extent of customer satisfaction.

### B. Customer Needs and Expectations

The achievement of a strong customer satisfaction is closely related to the understanding customer needs and expectations (William and Bertsch, 1992). According to the Kano Model (2001), customer needs can be divided into:

- **Basic needs** – obvious needs of customers and if not met, he is dissatisfied, however

meeting this needs may not be enough for customer satisfaction. Its satisfaction results in "*must be quality*".

- **Expected needs** – these are important needs that customers are fully aware of and satisfaction is expected in every purchase; their satisfaction creates "*expected quality*".
- **Excitement needs** – these are unconscious and unspoken needs of customers. By identifying and satisfying such needs, companies will have added large value to customers and can win loyal customers. This satisfaction creates "*attractive quality*".

### C. Perceived Value

Perceived value is defined as "the results or benefits customers receive in relation to total costs (which include the price paid plus other costs associated with the purchase) or the consumers' overall assessment of what is received relative to what is given" (Holbrook, 1994 and Zeithaml, 1988).

Additionally, Zeithaml (1988) found out that customers who perceive that they receive value for money are more satisfied than customers who do not perceive they receive value for money.

Past research studies suggested that there are four features, which are key drivers of the customer value of cellular services: network quality, price, customer care, and personal benefits (Booz, Allen & Hamilton, 1995, Danaher & Rust, 1996; Bolton, 1998; Gerpott, 1998; Wilfert, 1999).

- The network quality refers to excellent indoor and outdoor coverage, voice clarity, and no connection breakdowns.
- Price refers to what is paid to obtain access to use the network.
- Customer care refers to the quality of the information exchanged between customer and supplier or network provider in response to enquiries and other activities initiated by the network provider, for example presentation of invoices.

- Personal benefits refer to the level of perception of the benefits of mobile communications services by individual customers.

#### D. Service Quality

Another factor that contributes to satisfaction is service quality. Service quality is defined as “the difference between customer expectations and perceptions of service” or “as the customers’ satisfaction or dissatisfaction formed by their experience of purchase and use of the service” (Gronroos, 1984 and Parasuraman et al.1988).

Previous studies on mobile telecommunication services, measured services quality by call quality, pricing structure, mobile devices, value-added services, convenience in procedures, and customer support (Kim, 2000; Gerpott et al., 2001; Lee, Lee, & Freick, 2001).

#### E. Internal Satisfaction

Research works have shown the importance and the link of internal (employee) satisfaction to the external (customer) satisfaction. Hill and Alexander (2000) stated that there is a positive relationship between employee satisfaction and customer satisfaction and this is achieved in companies that practice employee motivation and loyalty. They reported that “*employees that are more motivated to achieve customer satisfaction tend to be more flexible in their approach to their work, make fewer mistakes and use more initiative*”.

#### F. Complaint Management

Albrecht and Zemke (1985) found that of the customers who register complaints, between 54% and 70% will do business again with the company if their complaints are resolved. This figure increases to 95% if the customer feels that the complaint was resolved promptly. Customers who have complained to a company and had their complaints satisfactorily resolved tell an average of five people about the good treatment they received. Hart, et al.,(1990), reported that when the service provider accepts responsibility and resolves the problem when customers complain, the customer becomes “bonded” to the company.

Levesque and McDougall (1996) in their case study on retail banking found out that if a service problem or customer complaint is ill or not properly handled, it has a substantial impact on the customer’s attitude towards the service provider. However, the study did not support the notion that good customer complaint management leads to increased customer satisfaction. They reported that “*at best, satisfactory problem recovery leads to the same level of customer satisfaction as if a problem had not occurred*”.

### III. CONCEPTUAL MODEL OF THE STUDY

As it was found during the literature review that factors like, consumer needs and expectations, perceived value, service quality, internal satisfaction and complaint management are the factors which lead to customer satisfaction. Based on the above literature review researchers’ derived following conceptual model of the study as shown in figure 1.

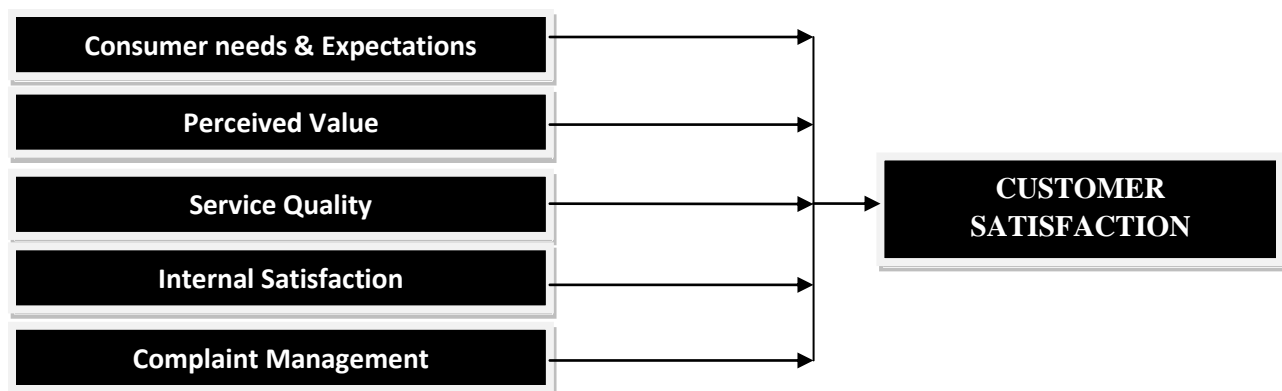


Figure 1: Dimensions of Customer Satisfaction in Services (Conceptual Model of the study)



#### IV. HYPOTHESES OF THE STUDY

Looking at the above conceptual model of the study, researcher has derived following hypotheses for the current study:

**H<sub>0\_1</sub>:** There is no relationship between network quality and customer satisfaction with GSM service provider.

**H<sub>0\_2</sub>:** There is no relationship between billing and customer satisfaction with GSM service provider.

**H<sub>0\_3</sub>:** There is no relationship between validity period and customer satisfaction with GSM service provider.

**H<sub>0\_4</sub>:** There is no relationship between responsiveness support and customer satisfaction with GSM service provider.

**H<sub>0\_5</sub>:** There is no relationship of Attitude of Respondent and satisfaction with GSM service provider.

**H<sub>0\_6</sub>:** There is no relationship of Fulfillment and satisfaction with GSM service provider.

#### V. DATA ANALYSIS & INTERPRETATIONS

In order to test the relationship between the predefined factors of GSM services and customer satisfaction we will use Pearson's correlation technique as given below.

The correlation table 1 shows all the pair wise correlations. The values in the correlation table are standardized, and range from 0 to 1 (positive and negative). We observe that all variables are highly correlated (ranging from 0.308 to 0.871) with satisfaction. This means we have chosen a fairly good set of independent variables (Network, Billing, Validity, Responsiveness, Attitude, and Fulfillment) to try and correlate with satisfaction. These correlations are one-to-one correlations of each variable with others.

Looking at significance values given in the above correlation table, here we reject all null hypotheses, H<sub>0\_1</sub> to H<sub>0\_6</sub>, and state that all independent variables like Network, Billing, Validity, Responsiveness, Attitude, and Fulfillment are positively related with the satisfaction of customers.

The correlation table shows that independent variables are highly correlated with each other. This indicates that they are not independent of each other and only one or two can be used to predict the dependent variable (Satisfaction). Regression is helpful in eliminating some of the independent variables as all of them might not required to drive customer satisfaction with GSM services.

Table 1:  
Correlation among various dimensions of GSM Services and customer satisfaction

		<b>Network</b>	<b>Billing</b>	<b>Validity</b>	<b>Responsiveness</b>	<b>Attitude</b>	<b>Fulfillment</b>	<b>Satisfaction</b>
<b>Network</b>	<b>P</b>	1	.445**	.573**	.497**	.532**	.482**	<b>.527**</b>
	<b>Sig.</b>		.000	.000	.000	.000	.000	.000
<b>Billing</b>	<b>P</b>		1	.350**	.267**	.316**	.280**	<b>.308**</b>
	<b>Sig.</b>			.000	.000	.000	.000	.000
<b>Validity</b>	<b>P</b>			1	.665**	.704**	.771**	<b>.690**</b>
	<b>Sig.</b>				.000	.000	.000	.000
<b>Responsiveness</b>	<b>P</b>				1	.901**	.630**	<b>.871**</b>
	<b>Sig.</b>					.000	.000	.000
<b>Attitude</b>	<b>P</b>					1	.667**	<b>.961**</b>
	<b>Sig.</b>						.000	.000
<b>Fulfillments</b>	<b>P</b>						1	<b>.637**</b>
	<b>Sig.</b>							.000
<b>Satisfaction</b>	<b>P</b>							1

Table 2 :  
Regression Model of Customer Satisfaction.

Regression Model						
Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.317	0.104		3.056	.003
	<b>NETWORK</b>	0.229	0.062	0.205	3.703	<b>.000</b>
	<b>BILLING</b>	0.229	0.050	0.242	4.559	<b>.000</b>
	<b>VALIDITY</b>	0.084	0.055	0.083	1.534	.127
	<b>RESPONSIVENESS</b>	0.180	0.054	0.177	3.302	<b>.001</b>
	<b>ATTITUTDE</b>	0.009	0.021	0.023	0.441	.659
	<b>FULLFILMENT</b>	0.396	0.058	0.390	6.826	<b>.000</b>
a.) Dependent Variable: SATISFACTION,		b.) R Square: 0.924				

The regression model of the following (as shown in table2) form has been used by entering all the six 'X' variable in the model

$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6$   
and the value of a, b1, b2, b3, b4, b5, b6.

In the output of regression model the value of 'B' gives all the coefficients of the model which are as follow.

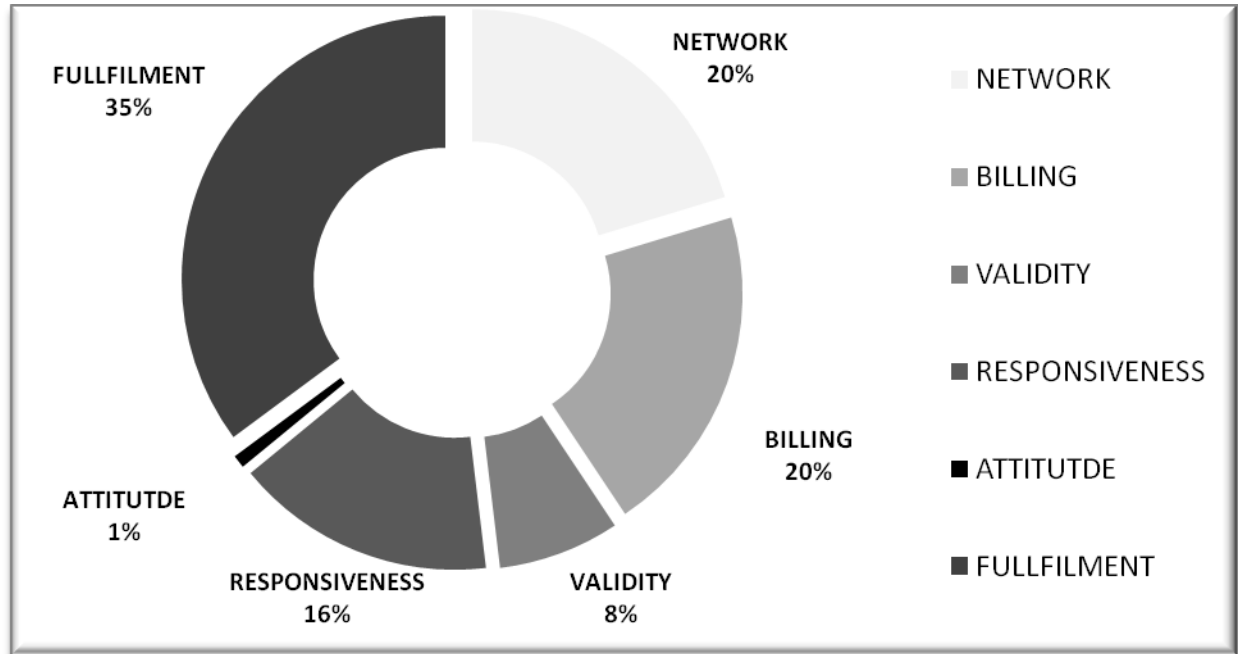
$X_1 = 0.229$ ,  $X_2 = 0.229$ ,  $X_3 = 0.084$ ,  $X_4 = 0.180$ ,  $X_5 = 0.009$ ,  $X_6 = 0.396$ , A (constant) = 0.317

These values can be substituted in the above equation to get the value of Y (satisfaction):

$Satisfaction = 0.317 + 0.229 (\text{Network}) + 0.229 (\text{Billing}) + 0.084 (\text{Validity}) + 0.180 (\text{Responsiveness}) + 0.009 (\text{Attitude}) + 0.396 (\text{Fulfillment})$ .

The p-level is observed to 0.000, indicating that the model is statistically significant at a confidence level of 99.99.

The R Square value is 0.924. We also note that t-Test for significance of individual dependent variables indicates that at the significance level of 0.05 (confidence of 95%), which indicates higher power of this regression model. Again we can find that all variables except 'VALIDITY' and 'ATTITUDE' are statistically significant in the regression model to determine the dependent variable satisfaction. These variables are individually significant at 95% confidence limit.



The equation we have obtained, mentioned above, means that satisfaction of customer will increase in a state if we improve network, or if providing better billing facility, or if increase validity, or if improve customer-care. The estimated increase in satisfaction of customer for every unit increase or decrease in these variables is given by respective variables. That means, if the network quality is increased by 1 percent, satisfaction is estimated to increase by 0.229 percent, provided that all other variable are unchanged. Similarly, if responsiveness increases by 1 percent; satisfaction is expected to increase by 0.18 percent, provided all other variable remaining constant.

There is one coefficient, that of the VALIDITY and ATTITUDE variable, which doesn't make too much intuitive sense. If we improve these two facilities the satisfaction is estimated to increase by 0.084 and 0.09 percent only. But if we look at the individual variable t-Test, we find that coefficient of

the independent variable VALIDITY and ATTITUDE are statistically not significant (significance level 0.127 and 0.659). Therefore it is not to be used when interpreting regression, as it may lead to wrong conclusion. So we conclude that independent variables like NETWORK, BILLING, RESPONSIVENESS and FULLFILMENT are statistically significant at 95% confidence level since T is less than 0.05. Therefore one should look at the relationship of satisfaction with these independent variables.

We have now compared the current regression model with the conceptual model of the study, where we found that only factor "Attitude of employees (Internal Satisfaction)", does not directly leads to the Customer Satisfaction, whereas rest of all service factors does leads to dependent variable Customer Satisfaction significantly. So here we will derive the current regression model compared to the conceptual model as shown in figure 2:

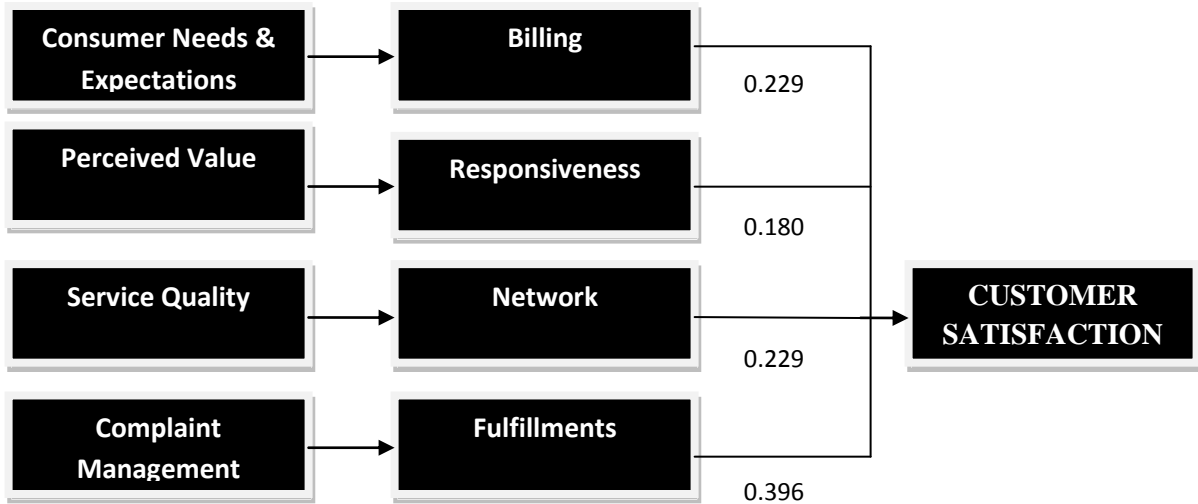


Figure 2: Dimensions of Customer Satisfaction in Services of GSM (Services Customer Satisfaction Model)

## VI. FINDINGS

The results demonstrated that most of the customers are satisfied with the performance of the mobile telecom operators. The implications of these findings are that network quality and complaint management are the most significant of all the mobile services attributes and its quality strongly affect satisfaction. Only Internal Satisfaction (Attitude) showed a weak correlation with satisfaction, because the employee's internal satisfaction with company may not directly influence customer's satisfaction. 100% of Airtel and Idea customer are satisfied, more than 80% of Vodafone customers are satisfied, but more than 40% of Reliance customers are dissatisfied. From multiple regression analysis, we can find that 8% validity, 16% responsiveness, 20% network, 20% billing, 35% fulfillment, influence the customer satisfaction.

Most of the customers are satisfied with the performance of the mobile telecoms industry in Ahmedabad. Based on the results of the data analysis, mobile operators need to maximize customer satisfaction to enhance successful business performance and an important way to achieve this is to improve on the mobile service attributes. Mobile operators can establish different market segments and adopt various strategies to satisfy needs of different

customers. The result also demonstrated that the combination of the mobile services attributes has strong influence on satisfaction.

## VII. RECOMMENDATIONS

- The problem is with the Network and the call centre services. Telecom service should work and try to make it efficient with more technical support and training programs.
- To expand the market the company should not only concentrate over the urban areas, but also have look at the untapped rural market. And establish good offers considering the villagers' requirements.
- Operators should try to win the trust of customers by providing better services and attractive schemes.
- Every customer wants satisfaction from the product so give him or her as much convenience as possible to make him or her loyal customer.
- Company has to improve attitude of their employee which is helpful to gain customer satisfaction indirectly.

## VIII. CONCLUSION

The implication of this finding is that network quality is the most significant of all the mobile services attributes and its quality strongly affect satisfaction. Billing, validity period and customer support showed weak influence on satisfaction. These results indicate that the evaluation of these factors without alignment is meaningless and have weak impact on satisfaction. Thus, to increase customer satisfaction, mobile operators should focus on improving mobile services attributes by investing in equipment to enhance call quality and coverage, offer reasonable pricing and price discounts, offer reasonable validity period and enhance customer care through routine personnel training and provision of better customer-friendly equipment.

## IX. LIMITATIONS

Survey is conducted by considering the time constraint. Sample size of two hundred customers may not be adequate. It is difficult to be precise about the most appropriate sample size for conducting such research. However, we expect follow-up studies to expand the sample range so that the relevant research might be more representative. Due to subjectivity and beliefs of consumers, some answers of the questionnaire may differ from the reality. Also due to lack of available literature on customer satisfaction of the mobile telephone industry and lack of cooperation from the respondents affect the validity of the data.

## REFERENCES

- [1] Adomi, E.E, 2005- Mobile telephony in Nigeria. Library Hi Tech News, Number 4 2005, pp. 18-21, Emerald Group Publishing Limited. The Electronic library.
- [2] Ahmad, J. and Kamal, N., 2002- Customer satisfaction and retail banking: an assessment of some the key antecedents of customer satisfaction in retail banking. International Journal of Bank Marketing, 20/4, pp.146-160.
- [3] Anderson, E.A. and Sullivan, M.W., 1993 – “The antecedents and consequences of customer satisfaction for firms”, Marketing Science, Vol.12 Spring, pp. 125-43.
- [4] Anderson, E.W., Fornell, C. and Lehmann, D.R. (1994), “Customer satisfaction, market share, and profitability: findings from Sweden”, Journal of Marketing, Vol. 58, July, pp. 53-66.
- [5] Athanassopoulos, A.D., 2000- “Customer satisfaction cues to support market segmentation and explain switching behavior”, Journal of Business Research, Vol.47 No.3, pp.191-207
- [6] Bhave, A., 2002- “Customer satisfaction measurement”, Quality and Productivity Journal, Symphony Technologies Pvt Ltd, Erendavane, India.
- [7] Bhote, K.R., (1996)- Beyond Customer Satisfaction to Customer Loyalty, AMA Management Briefing, New York, NY.
- [8] Bolton, R.N. (1998) - A dynamic model of the duration of the customer’s relationship with a continuous service provider: the role of satisfaction”, Marketing Science, Vol. 17 No. 1, pp. 45-65.
- [9] Eriksson, K and Löfmarck Vaghult, A., (2000)- “Customer retention, purchasing behavior and relationship substance in professional services”, Industrial Marketing Management, Vol.29 No.4, pp.363-72.
- [10] Fečiková, I., (2004) – An index for measurement of customer satisfaction. The TQM Magazine, Vol. 16, No 1, pp. 57-66.
- [11] Fornell, C. (1992) - A national customer satisfaction barometer: The Swedish experience. Journal of Marketing, 56, 6-21.
- [12] Gale, B., 1992- Monitoring Customer Satisfaction and Market Perceived Quality, Worth Repeating Series, No. 922CS01, American Marketing Association, Chicago, IL.
- [13] Gerpott, T.J., Rams, W., and Schindler, A., (2001) – “Customer retention, loyalty, and satisfaction in the German mobile telecommunications market, Telecommunications Policy, Vol. 25 No. 4, pp.249 – 269.
- [14] Gronhaug, K. (2002) – Research methods in Business Studies: A practical guide, 2<sup>nd</sup> ed., London: Prentice Hall.
- [15] Grönroos, C. (1984) – “A service quality model and its marketing implications”, European Journal of Marketing, Vol. 18 No. 4, pp. 36-44.
- [16] GSM Association (2006) - www.gsmworld.com
- [17] Hill, N. and Alexander, J. (2000) – Handbook of customer satisfaction and loyalty measurement, 2<sup>nd</sup> ed., Gower Publishing Ltd., England.
- [18] Hill, N., Brierley J., and MacDougall, R. (2003) - How to measure customer satisfaction, 2<sup>nd</sup> edition, Gower Publishing Ltd., England.
- [19] Hoff, Dean, (2006)- South African cellular wars in Nigeria. International Journal of Emerging Markets, Vol. 1 No. 1, 2006, pp. 84-95.
- [20] Homburg, C. and Giering, A. (2001) - “Personal characteristics as moderators of the relationship between customer satisfaction and loyalty – an empirical analysis”, Psychology and Marketing, Vol. 18 No. 1, pp. 43-66.
- [21] Hunt, H.K. (1982) - “A “10” based on expectations, but normatively a 3.6371”, in Day, R.L. and Hunt, H.K. (Eds), Proceedings of the 7<sup>th</sup> Annual Conference on Consumer Satisfaction, Dissatisfaction and Complaining Behaviour, University of Tennessee, Knoxville, TN, October, pp. 130-31.
- [22] Hunt, K.H. (Ed.) (1977)- “Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction”, Marketing Science Institute, Cambridge, MA.

- [23] Kano, N. (2001) – Life cycle and creation of attractive quality. In proceeding from Quality Management and Organization Development (QMOD) (Dahlgard, S.M Ed) Linkoping University, Sweden.
- [24] Keiningham, T.L. (2005)- Does customer satisfaction leads to profitability; share of wallet, as in Keiningham, T.L., Perkins-Munn, T. and Evans, H. (2003), “The impact of customer satisfaction on share-of-wallet in a business-to-business environment”, Managing Service Quality, www.emeraldinsight.com
- [25] Keiningham, T.L., Perkins-Munn, T. and Evans, H. (2003), “The impact of customer satisfaction on share-of-wallet in a business-to-business environment”, Journal of Service Research, Vol. 6 No. 1, pp. 37-50.
- [26] Keith, R.J. (1960) – “The marketing revolution”, Journal of Marketing”, Vol. 24, January, pp. 35-8.
- [27] Kim, H. (2000) - The analysis and determinants of customer loyalty in Korean mobile phone. Korean Information Society Review, 2000, 1–18.
- [28] Lee, J., Lee, J., & Freick, L. (2001) - The impact of switching costs on the customer satisfaction-loyalty link: Mobile phone service in France. Journal of Services Marketing, 15(1), 35–48.
- [29] McDougall, H.G., and Levesque, T (2000) – “Customer satisfaction with services: putting perceived value into the equation”, Journal of Services Marketing, Vol., 14 No. 5, pp. 392-410.
- [30] Naumann, E. (1995), Customer Satisfaction Measurement and Management: Using the Voice of the Customer, Thomson Executive Press, Cincinnati, OH.
- [31] Nelson, E., Rust, R.T., Zahorik, A., Rose, R.L., Batalden, P. and Siemanski, B., 1992- “Do patient perceptions of quality relate to hospital financial performance?” Journal of Healthcare Marketing, Dec., pp.1-13.
- [32] Niraj, R., Foster, G., Gupta, M. and Narasimhan, C. (2003)-“Understanding customer level profitability implications of satisfaction programs”, Washington University in St Louis Working paper, No. 2003-09-001, Olin School of Business, St Louis, MO.
- [33] Oliver, R.L. (1981) – “Measurement and evaluation of satisfaction process in retail setting”, Journal of Retailing, Vol. 57, pp. 25-48.
- [34] Oyewole, P., (2001)- Consumer's socio-demographic characteristics and satisfaction with services in the airline industry, services Marketing Quarterly, Binghamton, Vol.23, Iss. 2, pg.61
- [35] Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988) – “SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality”, Journal of Retailing, Vol. 64, Spring, pp. 2-40.
- [36] Parker, C., and Mathews, B.P., (2001) – “Customer satisfaction: contrasting academic and consumers' interpretations”, Marketing Intelligence and Planning, 19/1, pp. 38-44.
- [37] Pizam, A. and Ellis, T. (1999) – “Customer satisfaction and its measurement in hospitality enterprises”, International Journal of Contemporary Hospitality Management 11/7 [1999] 326-339.
- [38] Reichheld, F.F. (1996) - The Loyalty Effect, Harvard Business School Press, Boston, MA.
- [39] Rust, R.T. and Oliver, R.L. (1994) – “Service quality insights and managerial implications from the frontier”, in Rust, R.T. and Oliver, R.L. (Eds), Service Quality New Directions in Theory and Practice, Sage, CA
- [40] Taylor, S.A. and Baker, T.L. (1994) – “An assessment of the relationship between service quality and customer satisfaction in the formation of consumers' purchase intentions”, Journal of Marketing, Vol. 58, Summer, pp. 163-78.
- [41] Turel, O. and Serenko, A. (2006) - Satisfaction with mobile services in Canada: An empirical investigation, Telecommunications Policy 30 (2006), pp.314–331, 2006.
- [42] Wiele, T. Vander, Boselie, P. and Hesselink, M., (2002)- Empirical evidence for the relationship between customer satisfaction and business performance, Managing Service Quality, Vol. 12, No. 3, pp.184-93.
- [43] Yi, Y. (1990) – “A critical review of consumer satisfaction”, in Zeithaml, V.A. (Ed.), Review of Marketing, American Marketing Association, Chicago, IL, pp. 68-123.
- [44] [http://articles.timesofindia.indiatimes.com/2010-02-13/india-business/28116619\\_1\\_future-group-tata-tele-ties-t24](http://articles.timesofindia.indiatimes.com/2010-02-13/india-business/28116619_1_future-group-tata-tele-ties-t24)
- [45] [http://www.flinttelecomgroup.com/index.php?page=pressrelease&action=view&pressrelease\\_id=35](http://www.flinttelecomgroup.com/index.php?page=pressrelease&action=view&pressrelease_id=35)
- [46] <http://www.tmcnet.com/usubmit/2010/08/16/4958304.htm>

# CEO Remuneration in India: Identifying Relationship Between Pay And Performance.

A. Dr. Yagnesh M. Dalvadi, B. Shraddha G. Raj, c. Dharmendra K.Gohil

*Abstract*-Today, CEO compensation is most debatable and interesting topic in the world. It has generated a lot of debate among legislators, corporate observers, economists, journalists and management experts for payment of compensation to CEO. It has been attracted to investors, media, public as well as academia because of politics in pay of CEO compensation plays a very important role in financial performance of company. This paper examines the relationship of CEO compensation and firm's financial performance of selected companies. For this purpose we have selected eight index based companies. Correlation and regression has been used for testing of our hypothesis. The study found that the relationship between CEO compensation and company's performance have mixed relationship in selected company. Majority public sector companies' relationship was negative. However, majority private limited companies show positive relationship.

## 1. INTRODUCTION

The word compensation refers a monetary benefit given to employees in return for services provided by them. It includes basic salary, short-term incentives, bonus, long-term incentives, stock grants, stock options and other benefits. CEO compensation refers to the benefits paid to executive positions and the management team through the Board of Directors (BODs). It has been given by the company to its CEO for rendering his/her services. Remuneration to CEO is depends on company's performance. It means, it is depends on profitability of the company for the financial year. In different country, compensation paid to CEO is different. Some countries paid higher fraction of compensation in the form of short-term incentives while some are paid in the form of salary. In India, most of the companies paid higher fraction of compensation in the form of basic salary and commission. According to Companies Act 1956, compensation to CEO does not exceeding 11% of net profit for the financial year. But, most of the company does not follow this regulation and paid compensation to CEO on the basis of sales turnover. Hence, there is a great need to study this topic and find out the best results.

## 2. REVIEW OF LITERATURE

### 1. Scott J. Wallsten (2000)

This paper examines two main features of the relationship between executive compensation and firm performance. First, does the relationship itself change depending on firm performance? The author found that executives are rewarded in good years but not punished in bad years.

Second, does the relationship change with the executive's rank in the company? He found that the top executive's compensation is most strongly linked with performance i.e. on individual performance.

### 2. Yaron Amzaleg and Abraham Mehrez (2004)

This paper examines the statistically significant variables affecting the CEO's salary and compensation of top five executives. They used 186 public corporations, in which the CEO's total annual remuneration exceeded NIS-1 million in 1997. They also examine the policy regarding CEO's compensation and agency problem.

### 3. Yungsan Kim (2004)

This article examines the dynamics of chief executives turnover and its relation with firm performance. For this purpose he has been used a duration model which is incorporating up to ten years of firm performance. He found that the firm performance has a persistent effect on the chief executive's future turn over, except the performance in the early years of executive's tenure.

### 4. Ivan E. Brick, et. Al ( 2005)

They found a significant positive relationship between CEO and director compensation. They also find evidence that excess compensation is associated with firm under performance.

### 5. Mikko Makinen ( 2005)

This paper examines how CEO pay is related to firm size and to firm performance in Finland. For this purpose he has been used a new individual- level compensation data over the period from 1996 to 2002. They found that CEO average compensation has increased substantially over the period, especially compared to that of industrial worker average compensation.

### 6. Mark Farmer ( 2008)

This research paper discusses the measurement of the dependent chief executive compensation variable and investigates the relationship between pay and performance. They propose a methodological frame work for measuring chief executive compensation.

### 7. Melih Mandonoglu and Ersem Karadag (2008)

This research paper examines the pay- for- performance sensitivity in the U.S. restaurant industry. This study also focuses on how CEO compensation changes with the firm performance.

**3. RESEARCH METHODOLOGY**

This study is based on secondary data collected from annual reports. Convenient sampling technique has been used for selection of samples. Total eight companies viz. BHEL, L&T, ONGC, Reliance, Hindalco, Tata Steel, NTPC and Tata Power have been selected which are parts of BSE Sensex index. The researcher has selected those Companies which are top in market cap. The market cap value is based on 28th October, 2010 referring from Economic Times. The data was taken for the period 2005-06 to 2009-10 for all sample firms.

**4. OBJECTIVES OF THE STUDY**

Following are objectives of the study

1. To study the CEO compensation/remuneration of selected companies.
2. To compare the pay and performance of selected companies.
3. To find out whether CEO’s compensation is paid according to firm performance.

**5. HYPOTHESIS**

1. There is no impact of CEO’s compensation on profitability of selected companies.
2. There is no significant different of correlation between profitability ratios and total compensation of selected companies.

Appropriate statistical techniques were used for testing the hypotheses.

**6. CEO COMPENSATION OF SELECTED COMPANIES**

Table: 1 CEO Compensation of BHEL (Rupees in Lakhs)

Year	Salary	Benefits	Performance Linked Incentives	Total
2005-06	31.60	35.20	1.20	67.90
2006-07	37.90	40.30	2.40	80.70
2007-08	79.90	50.90	7.60	138.40
2008-09	42.20	30.20	5.90	78.30
2009-10	69.90	113.50	7.70	191.10
Average	52.30	54.00	5.00	111.30

The table no. 1 shows CEO compensation of BHEL from the year 2005-06 to 2009-10. highest fraction of compensation paid in the form of salary and benefits.

Table: 2 CEO Compensation of L&T (Rupees in Lakhs)

Year	Salary	Perquisites	Retirement Benefits	Commission	Total
2005-06	239.70	171.10	328.80	978.10	1717.80
2006-	540.0	281.80	518.30	1379.50	2719.5

07	0				0
2007-08	567.00	292.80	746.10	2196.40	3802.30
2008-09	594.00	386.80	1118.10	3546.90	5645.80
2009-10	621.60	450.00	1363.80	4429.70	6865.10
Average	512.50	316.50	815.00	2506.10	4150.10

The above table depicts that L&T Company paid more compensation in the form of commission which is based on net profit for the financial year.

Table: 3 CEO Compensation of ONGC (Rupees in Lakhs)

Year	Salary including DA	Other Benefits & Perks	Performance Incentives	Contribution to P.F .etc.	Provision for Leave s Gratuity	Total
2005-06	37.30	11.40	4.40	4.70	0.00	57.80
2006-07	38.90	31.70	14.10	5.80	0.00	90.60
2007-08	48.00	21.50	10.40	5.70	0.00	85.70
2008-09	78.30	35.20	17.40	7.90	28.20	167.10
2009-10	163.00	26.20	134.60	14.60	5.10	343.50
Average	73.10	25.20	36.20	7.70	6.70	148.90

From the above table, we can say that CEO of ONGC Company receives more salary than other benefits.

Table: 4 CEO Compensation of Reliance (Rupees in Lakhs)

Year	Salary	Perquisites & Allowances	Commission	Retiral Benefits	Total
2005-06	141.00	160.00	4023.00	0.00	4324.00
2006-07	128.00	150.00	4359.00	0.00	4637.00
2007-08	131.67	163.09	6413.16	45.14	6753.06
2008-09	134.00	166.00	3423.00	98.00	3821.00
2009-10	742.00	557.00	1994.00	797.00	4090.00
Average	255.334	239.218	4042.432	188.028	4725.01



It can be inferred that the CEOs of Reliance Company received highest portion in the form of commission which is based on net profit for the financial year of the Company.

Table: 5 CEO Compensation of Tata Steel

Year	Salary	Perquisites & Allowances	Commission	Total
2005-06	163.10	91.04	275.00	529.14
2006-07	152.00	60.05	375.00	587.05
2007-08	130.53	98.29	455.00	683.82
2008-09	88.00	69.21	350.00	507.21
2009-10	134.40	67.69	500.00	702.09
Average	133.606	77.256	391	601.862

The above table shows CEO compensation of Tata Steel Company. We found that the CEOs of the company received highest fraction of compensation in the form of commission.

Table: 6 CEO Compensation of Hindalco

Year	Salary, Benefits, Bonuses, Pension
2005-06	378.09
2006-07	494.00
2007-08	823.66
2008-09	1109.40
2009-10	1315.14
Average	824.058

Hindalco has only one executive officer. So, the company does not give a detail of each component of compensation. From the above table, we can say that the compensation paid to CEO is in increasing order year by year.

Table: 7 CEO Compensation of NTPC (Rupees in Lakhs.)

Year	Salary	Benefits	Bonus/Commission	Performance Incentives	Total
2005-06	66.38	18.58	0	9.80	94.76
2006-07	49.38	9.78	0	7.64	66.80
2007-08	61.42	23.38	0	33.51	118.31
2008-09	89.35	27.03	0	24.93	141.31
2009-10	136.11	67.15	0	58.86	262.12
Average	80.53	29.18	0	26.95	136.66

NTPC Company pays highest portion of compensation in the form of salary. The company does not paid bonus/commission to CEOs.

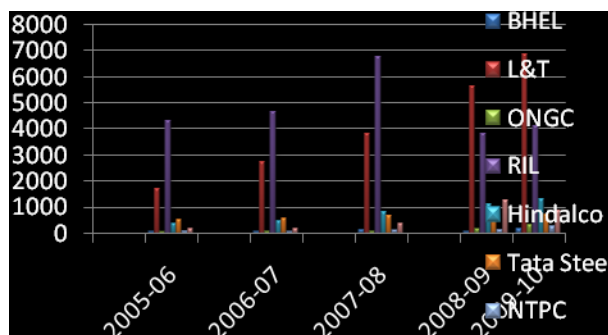
Table: 8 CEO Compensation of Tata Power (Rupees in Lakhs.)

Year	Salary	Commission	Perquisites	Retirement Benefits	Total
2005-06	37.20	62.50	54.15	34.42	188.27
2006-07	48.23	45.00	93.51	10.06	196.80
2007-08	91.04	135.00	115.77	21.70	363.51
2008-09	840.00	195.00	197.01	36.94	1268.95
2009-10	152.40	446.00	248.36	41.15	887.91
Average	233.774	176.7	141.76	28.854	581.088

From the above table, we found that CEOs of Tata Power gets more salary compare to other part of compensation.

TABLE: 9 CEO COMPENSATION OF SELECTED COMPANIES (Rs. In lacs)

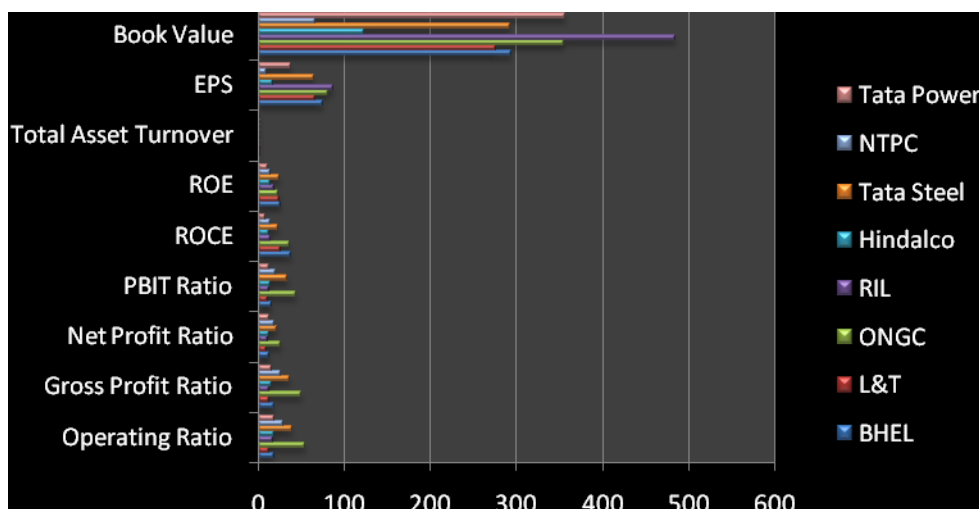
Year	BHEL	L&T	ONGC	RIL	Hindalco	Tata Steel	NTPC	Tata Power
2005-06	67.9	1717.8	57.8	4324	378.09	529.14	94.76	188.27
2006-07	80.7	2719.5	90.6	4637	494	587.05	66.8	196.8
2007-08	138.4	3802.3	85.7	6753.06	823.66	683.82	118.31	363.51
2008-09	78.3	5645.8	167.1	3821	1109.4	507.21	141.31	1268.95
2009-10	191.1	6865.1	343.5	4090	1315.14	702.09	262.12	887.91
Average	111.28	4150.1	148.94	4725.012	824.058	601.862	136.66	581.088



The above table and chart shows total compensation of selected companies for the study period. The highest compensation was paid by RIL in the year 207-08 and the lowest was paid by ONGC in the year 2005-06. When we look from overall performance, CEOs of RIL has got a highest compensation and CEOs of BHEL has paid lower compensation compare to other selected companies.

TABLE: 10 PROFITABILITY RATIOS AND OTHER RATIOS OF SELECTED COMPANIES

Average Ratio	BHEL	L&T	ONGC	RIL	Hindalco	Tata Steel	NTPC	Tata Power
Operating Ratio	17.97	11.83	53.62	17.1	18.12	38.76	28.5	17.99
Gross Profit Ratio	17.86	11.88	50.14	12.85	16.14	36.28	26.36	15.11
Net Profit Ratio	12.7	8.92	26.1	11.04	12.83	22.16	18.79	13.14
PBIT Ratio	15.92	10.72	43.81	12.75	14.26	33.68	20.57	11.9
ROCE	38.41	25.61	35.94	14.67	12.47	23.32	13.36	8.1
ROE	25.67	24.23	23.18	18.05	14.13	24.39	13.67	10.65
Total Asset Turnover	1.93	1.98	0.69	1.1	0.72	0.65	0.44	0.51
EPS	75.57	65.93	81.24	86.31	16.57	65.2	8.98	37.41
Book Value	293.39	275.97	354.86	485.3	123.3	293.19	65.62	356.3

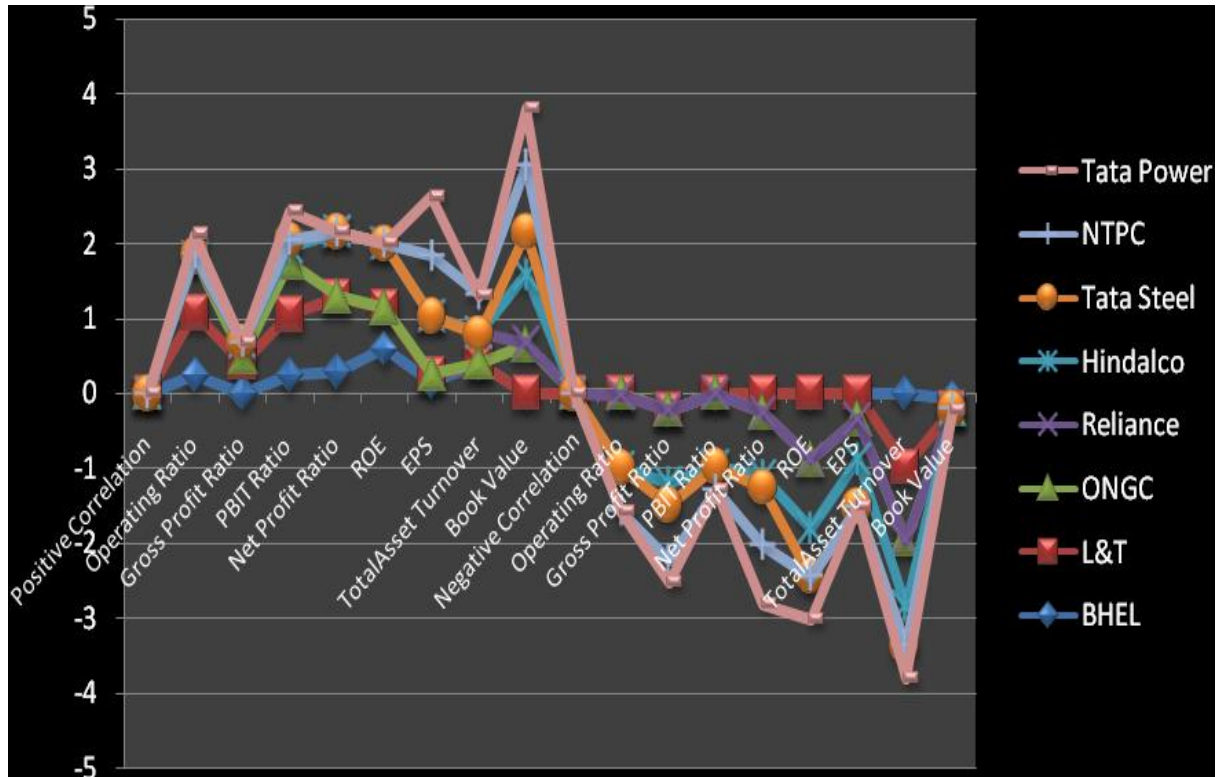


There are two words i) profit and ii) ability. Profit is a difference between incomes over expenses, while ability means capacity to earn profit. So, profitability means an ability to earn profit for operating a business. The term profitability implies the profit-making ability of business enterprise. The above table shows the profitability ratios of selected companies. Operating ratio, Gross profit ratio, Net profit ratio, PBIT ratio of ONGC is highest and L&T has lowest compared to other selected companies. ROE of BHEL is higher i.e.25.67 and Tata Power has lower i.e. 10.65. RIL has highest EPS and Book Value than others and NTPC has lower.

Correlation test is used to find out correlation between profitability ratio and CEO compensation of selected companies during the study period. Following is correlation value.

RELATIONSHIP BETWEEN CEO REMUNERATION ON FIRM PERFORMANCE

Ratios	BHEL	L&T	ONGC	Reliance	Hindalco	Tata Steel	NTPC	Tata Power
<b>Positive Correlation</b>								
Operating Ratio	0.250	0.825	0.735			0.043		0.278
Gross Profit Ratio		0.405	0.075	0.198				
PBIT Ratio	0.243	0.810	0.690	0.210		0.093		0.398
Net Profit Ratio	0.288	0.997		0.874				
ROE	0.595	0.544		0.854			0.013	
EPS	0.142	0.111		0.781			0.797	0.798
TotalAsset Turnover	0.405			0.394			0.515	
Book Value			0.648	0.034	0.902	0.576	0.871	0.779
<b>Negative Correlation</b>								
Operating Ratio				(0.014)	(0.984)		(0.571)	
Gross Profit Ratio	(0.227)				(0.971)	(0.289)	(0.736)	(0.285)
PBIT Ratio					(0.970)		(0.313)	
Net Profit Ratio			(0.254)		(0.854)	(0.145)	(0.770)	(0.776)
ROE			(0.887)		(0.923)	(0.631)		(0.575)
EPS			(0.343)		(0.541)	(0.625)		
TotalAsset Turnover		(0.979)	(0.978)		(0.890)	(0.524)		(0.428)
Book Value	(0.063)	(0.147)						



The above table and chart shows the correlation of CEO compensation and profitability ratios of selected eight companies. From this we can say that the linear correlation of Reliance, BHEL and L&T have strong positive result while Hindalco Company has strong negative correlation. It means there is opposite relation between firm's profitability performance and CEO compensation of Hindalco Company.

**RESEARCH FINDINGS**

1. CEO of BHEL gets more remuneration by the way of salary and benefits which are fixed in nature. However, performance linked incentives form a little part of remuneration. L&T Company gives compensation in the form of commission from net profit for the financial year. However, CEOs gets more benefits from company's earning. Reliance Industries decided remuneration in the same manner as of L&T Company. i.e. giving commission on net profit means more profit-more commission and less profit-less commission. Hence, CEOs are receiving higher compensation in the form of commission and benefits. CEOs of Tata steel get more compensation by way of commission. However, salary is fixed and this is not a main motto of the company, but to give the commission and CEOs gets motivated to work more. Hindalco has only one executive director receiving compensation under scheduled XIII of the Companies Act 1956, which is based on target oriented project. In NTPC company, CEOs getting higher fraction of compensation in the form of salary and benefits. However, when

company introduces new performance scheme, they are obtaining more performance based incentives.

2. CEO of public sector companies gets more fixed component of compensation. However, private sector companies' CEO gets more compensation as a variable component which is based on performance.

3. Statically results shows that there is significant difference in Operating Profit Ratio, Gross Profit Ratio, PBIT Ratio, Net Profit Ratio, ROCE, ROE, Total Asst Turnover, EPS and Book Value of selected companies during study period.

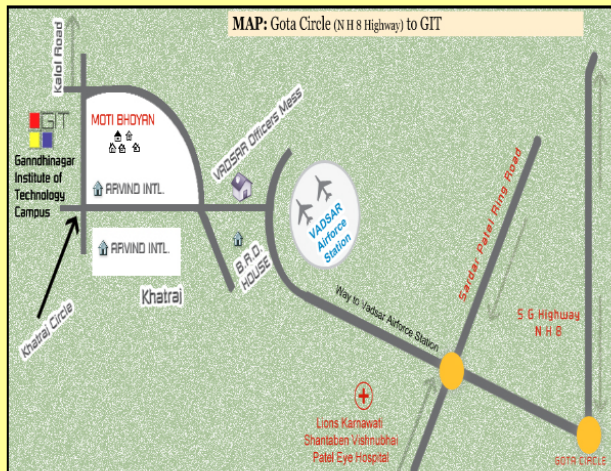
4. Correlation of BHEL, L&T and Reliance found more positive relationship compare to other selected companies, while Hindalco has strong negative correlation. From the correlation, we found that there is a positive relationship between firm's profitability performance and compensation paid to CEOs of BHEL, L&T and Reliance. It means it is paid according to companies act, if profit of a company is increased than the compensation of CEOs will increased. Hindalco Company has only one executive officer and when we look its linear correlation we can say that it has strong negative relation. It means the compensation of CEO is increased year by year while its profit shows decreasing trend. It is found that in BHEL, L&T, Reliance compensation to CEOs is paid according to firm performance, result of ONGC, NTPC and Tata Power shows middle result and Hindalco Company does not paid according to firm performance.

## REFERENCES

1. Chandra, P. (2008), Financial Management Theory and Practice, 6<sup>th</sup> Edition, 8.  
Tata McGraw Hill Education Private Limited, New Delhi.
  2. Kothari C.R., Research Methodology, Methods and Techniques, Second 9.  
Revised Edition (2004), New Age International Publishers, New Delhi.
  3. Brick Ivan E., Palmon Oded, and Wald John K. (2005), CEO Compensation, 10.  
Director Compensation, and Firm Performance Evidence of Cronyism? JCF  
Special Issue on Corporate Governance.
  4. Farmer Mark (2008), Chief Executive Compensation and Company 11.  
Performance: A Weak Relationship or Measurement Weaknesses? 17<sup>th</sup> 12.  
EDAMBA Summer Academy. <http://ssrn.com/abstract=1114647>.
  5. Kim Yungsan (2004), Long-Term Firm Performance and Chief Executive 13.  
Turnover: An Empirical Study of the Dynamics, the Journal of Law 14.  
Economics, 4 Organizations, V12 N2. 15.
  6. Madonoglu Melih and Karadag Ersem (2008), Firm Performance and CEO 16.  
Compensation: Reflections from the U.S. Restaurant Industry, Proceeding of 17.  
2008 EABR & TLC Conferences, p.p. 480-496, Rothenberg, Germany. 18.
  7. Mikko Makinen (2005), CEO Compensation, Firm Size and Firm 1.  
Performance: Evidence from Finnish Panel Data, Paper provided by The  
Research Institute of the Finnish Economy in its series Discussion Papers with  
number 1084.
  - Murphy, Kevin J., Executive Compensation (1998). Available at SSRN:  
<http://ssrn.com/abstract=163914> or doi:10.2139/ssrn.163914.
  - Wallsten Scott J. (March2000), "Executive Compensation and Firm  
Performance: Big Carrot, Small Stick." SIEPR Working Paper 99-17.
  - Yaron Amzaleg, Abraham Mehrez , The 'One Million Club: Executive  
Compensation and Firm Performance, \*Israel Economic Review Vol. 2, No.  
1, 107–147.
- Annual Reports
1. Annual Report of BHEL for the study period 2005-06 to 2009-10
  2. Annual Report of Hindalco for the study period 2005-06 to 2009-10.
  3. Annual Report of L&T for the study period 2005-06 to 2009-10
  4. Annual Report of NTPC for the study period 2005-06 to 2009-10
  5. Annual Report of ONGC for the study period 2005-06 to 2009-10
  6. Annual Report of Reliance for the study period 2005-06 to 2009-10
  7. Annual Report of Tata Power for the study period 2005-06 to 2009-10
  8. Annual Report of Tata Steel for the study period 2005-06 to 2009-10.
- News Papers
1. Economics Times of 28<sup>th</sup> October, 2010.
  2. Reform CEO Pay, Times of India



<a href="#">Home</a>	<a href="#">Trustee</a>	<a href="#">Editorial Board</a>	<a href="#">Director Message</a>	<a href="#">Papers</a>	<a href="#">Contact Us</a>
----------------------	-------------------------	---------------------------------	----------------------------------	------------------------	----------------------------



#### **Editorial Chief**

Dr. N M Bhatt

#### **Co-Ordinators**

Prof. Mitesh Mungla  
Prof. Svapnil Vakharia  
Prof. Gunjan Jani

#### **Institute Address**

Gandhinagar Institute of Technology  
Khatraj - Kalol Road,  
Village Moti Bhoyn,  
Tal. Kalol  
Gandhinagar-382721

Telephone No: +91-02764-281860/61

Fax No: +91-02764-281862

E-mail Address: [director@git.org.in](mailto:director@git.org.in)

#### **Office Address:**

Platinum Foundation Trust  
A-201, 202, Iscon Park,  
Opp, Star India Bazar,  
Satellite, Ahmedabad  
Gujarat, INDIA.

Telephone No: +91-079-26922627

Fax No: +91-079-26922628

E-mail Address: [trustee@git.org.in](mailto:trustee@git.org.in)