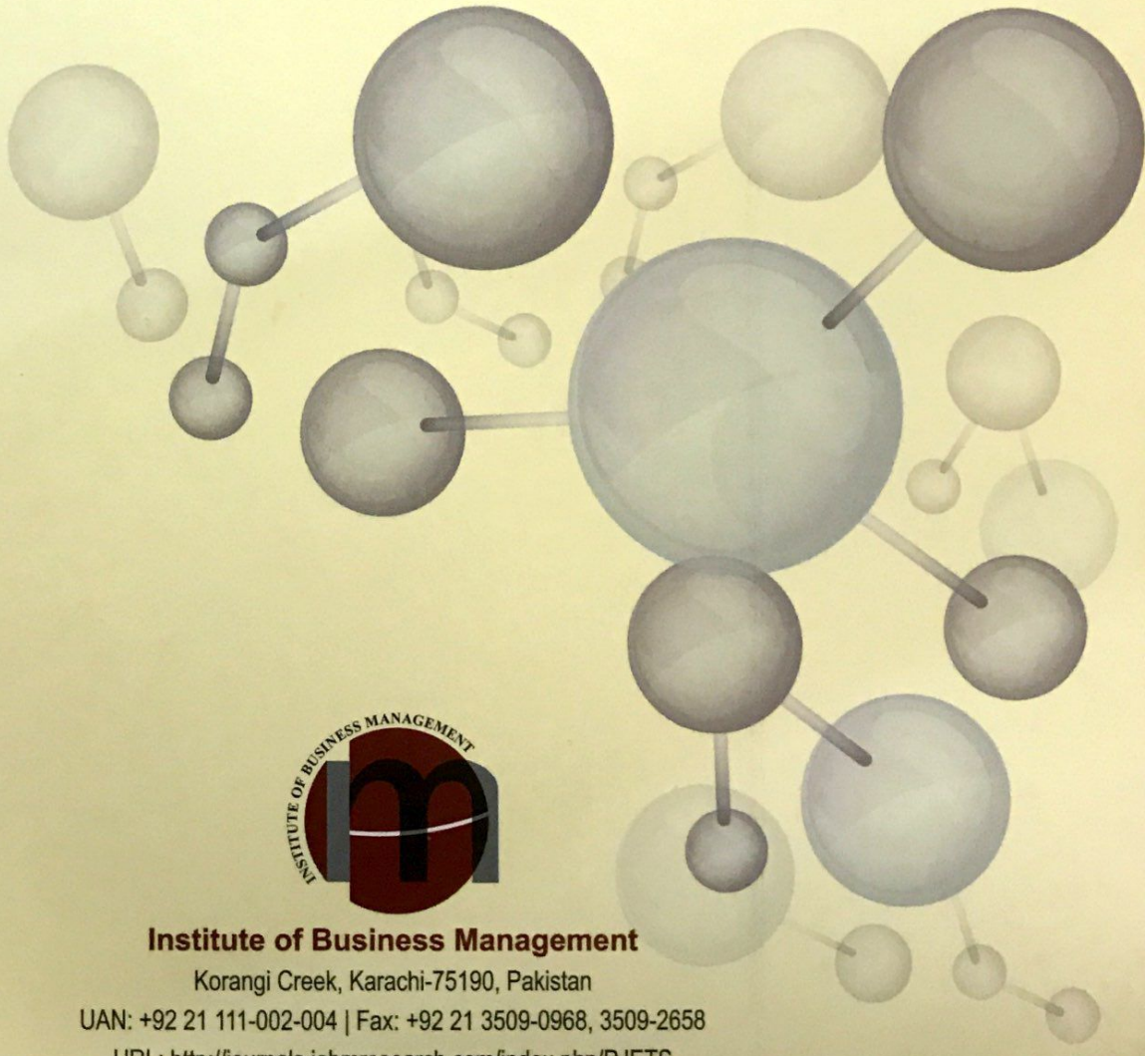
ISSN 2224-2333 (Online)
ISSN 2222-9930 (Print)

PAKISTAN JOURNAL OF ENGINEERING TECHNOLOGY AND SCIENCE

HEC APPROVED "Z" CATEGORY RESEARCH JOURNAL

VOLUME 6 NUMBER 2

DECEMBER 2016



Institute of Business Management
Korangi Creek, Karachi-75190, Pakistan
UAN: +92 21 111-002-004 | Fax: +92 21 3509-0968, 3509-2658
URL: http://journals.iobmresearch.com/index.php/PJETS
URL: http://www.iobm.edu.pk | e-mail: pjets@iobm.edu.pk

Editors' Note

We have completed six years of publication. Our mission was to provide a platform to the researchers, faculty and students to spread their findings. The goal was to link people from different ways of life to encourage others research contribution. This was not an easy task. Gathering researchers from different fields and convincing them to work as editorial board member, was a tedious job. We fortunately succeeded in developing an editorial board comprising of reputed scholars from different parts of the globe. Researchers from different Universities and research organizations were kind enough to submit their articles for possible publication. Internal and external reviewers worked with dedication, going through each submitted article, preparing and giving feedback to improve the quality of the submitted articles, is appreciated.

Scientific development is at a much faster speed and we are reading about new innovations from almost everywhere. We as a nation have to struggle to popularize the innovation culture. Drive out fear and let our scholar be bold enough to ask questions and to develop solutions.

University curriculum should be improved to allow young researchers to come up with new ideas, popularize proposed procedures, develop new procedures and help users and students.

Note: Issue based on papers presented in the conference are not subject to the standard of PJETS.

Scope of the Journal

PJETS is publishing and promoting innovative ideas and original research in the field of Science, Technology and Engineering since 2011, twice a year. This journal aims at publishing authentic research papers to create a culture of innovation and scientific development. The scope of the journal is limited to "Computer Sciences", "Engineering", relevant "Emerging Technologies", "Mathematics" and "Statistics".

Member Editorial Review Board (International – Academia)

Member Editorial Review Board (International – Non Academia)

Member Editorial Review Board (National – Academia)

Member Editorial Review Board (National – Non Academia)

Contents

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

An overview of Customer Relationship Management Software in Business Organizations

[1]Wajahat Hussain, [2*]Maria Andleeb Siddiqui and [2]Najmi Ghani Haider
[1]Karachi Shipyard and Engineering Works Limited, Karachi, Pakistan
[2]Department of Computer Science and Information Technology, NED University, Karachi, Pakistan

***Abstract*- The aim of the article is to present details of the research on focus on the customer in relation to the utilization of the customer relationship management computerized system in business organization. Customer relationship management (CRM) plays a central role in the growth and success of corporations in an environment of fast technological development and the extensive competition currently in evidence. CRM empowers corporations with better customer awareness and helps to construct sustainable relationships with customers. People, technology, and processes are the three basic components of CRM. In order to ensure the successful implementation and adoption of the CRM initiative, this paper presents a detailed review of the literature relating to CRM processes and its computerized system. Various types and levels of the CRM process as well as an understanding of the different perspectives of CRM are presented in this review. The paper suggests that it is important for an organization to understand the four major perspectives of CRM processes i.e. customer facing level processes, customer oriented processes, cross functional CRM processes and CRM macro-level processes. A survey is also conducted in this paper and a chi square test is performed to test the hypothesis relating to the efficiency of customer relationship management and the client relationship it offers. Results show that the hypothesis is statistically significant.**

Keywords: Computerized, Customer Relationship Management (CRM), processes, successful, sustainable.

I. Introduction

The main concept of customer relationship management is that the better knowing customers enable business organizations to ensure that they remain loyal on a permanent basis and also make it possible for them to serve customers better. CRM, in brief, is still going through a growing process Business strategy, Business philosophy, Business process and Technological tools are the main elaborations of CRM as shown in Figure 1. By offering a more customized and responsive service to each customer, CRM aims to increase customer loyalty and satisfaction and behaves as a customer focused business strategy [2]. Process management creates a superior customer value and retention and behaves as a relationship oriented tool when CRM is considered as Business Philosophy [1]. When considered as a Business Process, CRM is used to create customer knowledge and prospect

*mandleeb@neduet.edu.pk

identification and behaves as a "macro-level" process that creates the above defined sub processes [3]. CRM is also a technology for an organization, designed to foster a closer relationship with customers [4]. Information Technology mediated CRM as the policy, philosophy and coordinating strategy. With the help of CRM two way communications between the customers is created. This helps organizations to have an intimate knowledge of their wants, buying, selling patterns and needs. CRM initiatives have significant knowledge. CRM initiatives gain significant benefits. Most large organizations are investing significantly in CRM initiatives. CRM initiatives are supported by large and expensive CRM software packages.

A broad set of applications and software design is covered in CRM software that helps businesses to manage customer interactions, data, marketing, customer support and also to access business information and automates sales. Vendor, partner relations and employees are also managed by CRM software. Many products are scaled to a business of any size. CRM's need to adapt to the rise of social media is examined [18]. Effective ways of using social media on house core areas of CRM are also discussed and it is concluded that CRM strategies should use social media in order to better address and recognize challenges. The implementation of the CRM process in Islamic banks is empirically examined [19]. Along with this, there is also a focus on CRM's measurement of customer based profit performance and the need to develop a standard structural equal model for predicting that performance. Theoretical examination of CRM agenda is carried out [20] in relation to those organizations that have dominated the business world in recent years. The effectiveness of CRM agenda in relation to an organization with respect to customer loyalty and satisfaction is also explored. The focus is on what is required to implement a competent CRM system. The review concluded that for a successful implementation of CRM, the pairing of specific factors is required [21]. A study was conducted [22] in which the factors that affect the deployment of CRM at Nestle Company were identified and ranked.

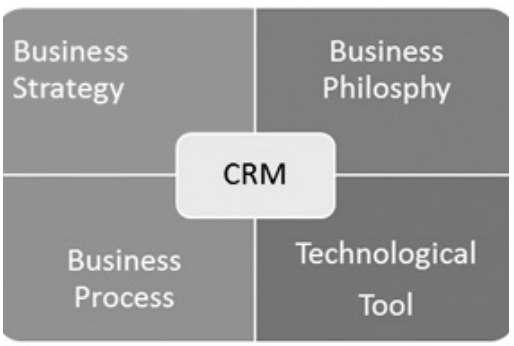


Figure 1 Elaborations of CRM

Figure 2 Levels of CRM Extension

II. Detail Overview of CRM

1. CRM LEVELS AND ITS TYPES/FORMS

CRM will act on an organization to the extent covered by CRM types/forms. There are four levels of CRM extension as shown in Figure 2 above.

A. Strategic

The creation of a customer-centric business culture is brought about by the strategic CRM. Through this culture a better value over competitors is created as it helps to determine where the resources of organizations can be better invested.

B. Operational

Customer's processes automation is brought about by an operational CRM. Automation includes sales-force automation, marketing and service automation. Automation and streaming workflow at the front office is dealt with by CRM. This includes processing of transactions, controlling of sales, marketing and services, workflows and data collection. [5][6][9][10][11].

C. Analytical

Using customer data, the increase in number of customers and organization value is dealt with by an analytical CRM. Operational CRM is the basis of the building of an analytical CRM which helps to analyze data of customers for information creation about customer segmentation, customer behavior and, most importantly, customer value obtained by using several statistical and data-mining tools. [5][10].

D. Cooperative

This can be the subset of an operational CRM because it works at a CRM operational level. Using a set of interaction channels, it focuses on customer integration [5][6][7][12][13] and selected business partners and customer suppliers working intimately. Different communication means such as emails, phone calls, website pages and fax are included in collaborative technologies which are used for interactions by the customers [14]. The involvement of CRM processes is mandatory when dealing with customer data.

2. CUSTOMER RELATIONSHIP MANAGEMENT PROCESSES

This section discusses how the CRM processes intercommunicate and integrate within the three levels of CRM kinds. The way in which things are done in the context of cooperation is defined as the process [8][15]. Such cooperation performs certain activities which are related to CRM processes and concern the management of the customer relationship with these activities being grouped together according to a longitudinal view of the relationship. The purpose of the

CRM process is to form customers' perceptions of cooperation and its products through identifying customers, making customer knowledge, and constructing customer relationships.

A. Categories of CRM Processes

The categories of CRM processes are shown in Figure 3.

Figure 3 Categories of CRM Processes

i. Vertical processes

The business function completely places within CRM certain processes such as customer acquisition process. These processes are termed as vertical.

ii. Horizontal processes

Cross functional processes such as product development are referred to as horizontal processes.

iii. Front-office processes

Customer-facing processes such as the complaint management process are alluded to as front office processes.

iv. Back-office processes

The non-facing and processes hidden from customers are referred to as back office process: for e.g. procurement processes.

v. Primary processes

Insurance cooperation has a co- directed process and courier cooperation has logistical processes. These processes are termed as primary processes. Usually these processes have vital revenue implications and cost.

vi. Secondary processes

The processes that have fewer cost or revenue implications for corporations are termed as secondary processes [16].

B. Themes of CRM processes

Figure 4 shows the four different views and perspective of CRM:

Figure 4 Views of CRM

a. Customer-facing level CRM processes

Management of the customer relationship initiation, termination and maintenance is a systematic process [16]. In order to maximize the value of the relationship portfolio management is done across all customer contact points.

There are three CRM processes at the customer-facing level of CRM:

- *Relationship initiation*
- *Relationship maintenance,*
- *Relationship termination* [16].

i. Relationship initiation

The activities that take place in the early stages or before the relationship, are referred to as initiation processes. These include identification of potential customers.

ii. Relationship maintenance

The activities by which normal customer relationships are portrayed are referred to as maintenance processes such as upselling, retention for organization and cross-selling.

iii. Relationship termination

The activities that help to resolve the problem of ending a bad relationship are such as are involved in termination processes. Termination management activities are connected to low-value customers. Unprofitable relationships that end are examples of relationship termination [16].

b. Customer-oriented CRM processes

The customer activities that are formed for need and satisfaction and act as a problem solver are encompassed in customer processes. The semi-structured and knowledge intensive nature of the

Customer oriented CRM processes.

There are three distinguished kinds of customer oriented CRM processes:

- *CRM delivery processes,*
- *CRM support processes,*
- *CRM analysis processes.*

i. CRM delivery processes

The processes of direct contact with customers are the CRM delivery processes. They are contemplated as part of the customer process inventory such as sales, campaign, complaint and service management processes.

i. CRM support processes

The processes used to deal for the accomplishment of supporting processes are the CRM support processes. It accomplishes this goal through loyalty management process and market research process.

ii. CRM analysis processes.

The process that concentrates on analyzing and combining the collected customer knowledge in other CRM processes is mentioned as CRM analysis processes. For example, feedback and knowledge management.

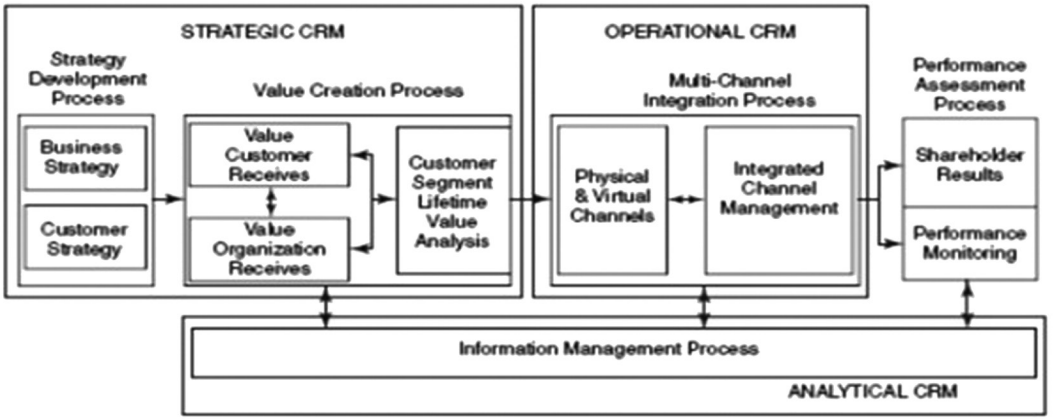


Figure 5. CRM forms and Processes

c. Cross-functional CRM processes

There are five generic cross-functional CRM processes which are based on a holistic approach of [17]:

- *the strategy development process*
- *the value creation process*

- *the multichannel integration process*
- *the information management process*
- *the performance assessment process.*

The CRM forms mentioned before and analytical, strategic and operations have the allocation of the CRM processes and how these processes inter-communicates with each other as shown in Figure 5. Four CRM processes involving strategic planning, information management, customer value, and performance measurement processes.

d. CRM Macro-level processes

The undertaken activities of cooperation that help to make market intelligence are referred to as macro-level processes. These processes help co-operations; through two sub processes i.e. knowledge management process and intercommunication management process, to sustain and construct a profit maximizing portfolios of customer relationships.

3. CRM PROCESS MODEL

In Figure illustrates the CRM process model. This model is demonstrated on the marketing and cooperate requirements for a successful CRM process.

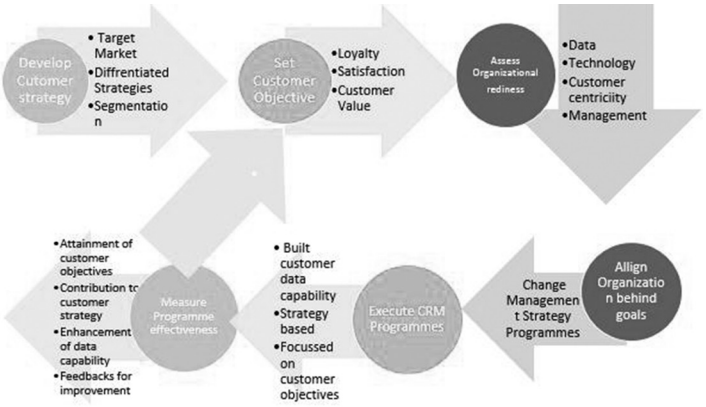


Figure 6 Illustration of CRM Process Model

The first step is the customer strategy development in which identification of target market is done. Development of differentiated strategies is done to deal with the customer's segments based on their profitability. For achieving customer satisfaction, value and loyalty, the setting of the customer objective is performed. After that the organization is focused on the customer rather than the product in order to access organizational readiness. Top management support and commitment are ensured for the planning of the CRM system implementation. In the next stage, the accomplishment of the enterprise wide program is done in such a way as to include changes made to the fundamental

processes of the organization or to transform the organizational culture i.e. product-centric to customer-centric. The above stage or step is needed to align the organization for CRM goals. After that specific programs help to focus on the differentiated strategies for all customer segments. In this way CRM programs are properly executed. The last stage involves measuring the effectiveness of CRM programs. For this purpose, a set of metrics is developed. This set is used to measure the effectiveness of CRM programs. For this purpose, a set of metrics is developed. This set is used for the measurement of certain metrics i.e. the contribution of each individual program to the overall CRM initiative. It is also designed to help measure the contribution of each program towards the enhancement of an organization's data capability and also to measure the potential for improving future programs. This model serves the purpose of illustrating the CRM process in simple terms. The model can, for the sake of better representation, can, of course, be improved and enhanced. A customer- oriented culture module can be added at the beginning. This helps to generate the need for and create an awareness of the importance of CRM. Another important aspect is the need to incorporate a feedback module for the purpose of developing an appropriate customer strategy. With the help of such a feedback module and on the basis of differentiated strategies and customer segmentation, the requisite target marketing can be accomplished.

III. CRM Software Based System

CRM software is designed to help businesses meet the overall goals of customer relationship management. Today's CRM software is highly measurable and customizable. It allows businesses to gain actionable customer awareness with a back-end analytical engine, enabling them to view business opportunities with the aid of predictive analytics and to streamline operations and privatize customer service on the basis of a customer awareness history and prior intercommunications with a given business. CRM software is commonly utilized to manage a business-customer relationship. However, CRM software systems are also utilized in the same way to manage business contacts, clients, contract wins and sales leads.

Cloud CRM is customer relationship management technology, where the CRM software, CRM tools and the cooperation's customer data resides in the cloud and is delivered to end-utilizers via the Internet. Cloud CRM typically offers retrieval to the application via Web-based tools logins where the CRM system administrator has previously defined retrieve levels over the cooperation. Employees can log in to the CRM system, simultaneously, from any Internet-allowed computer or device. Often, cloud CRM produces utilizers with mobile apps to make it easier to utilize the CRM on smartphones and tablets. Figure 7 shows the CRM dashboards.

Here, in the example of cloud CRM software from salesforce there are seven CRM software Functions classified in two kinds:

A. Sales & Marketing CRM Functionalities.

B. Customer Service CRM Functionalities.

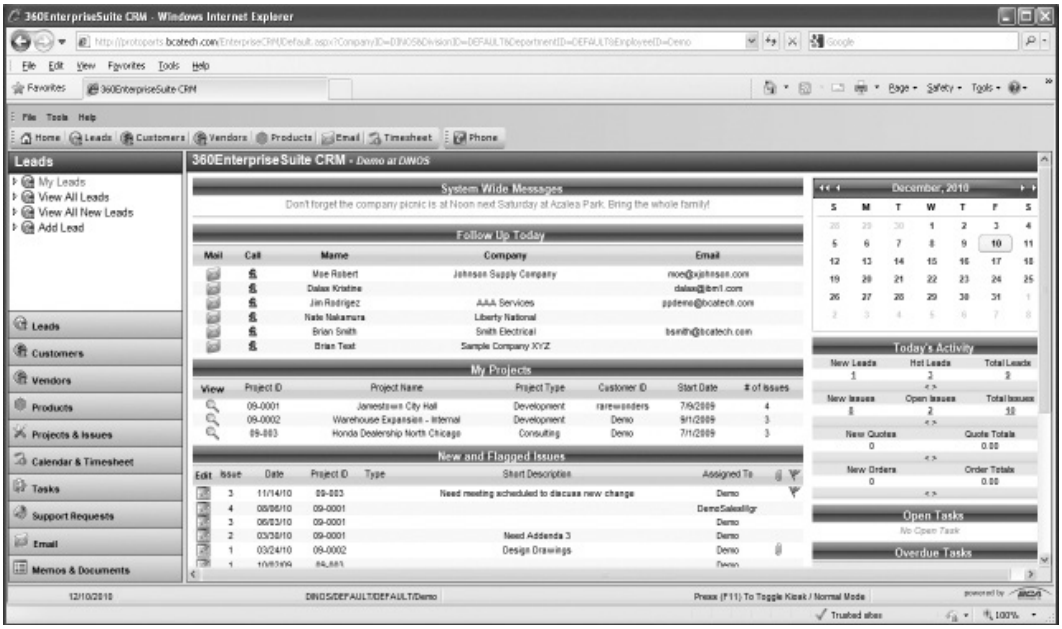


Figure 7 CRM dashboards

A. Sales & Marketing CRM Dashboard.

These dashboards are developed to relax people in different sales and marketing roles to measure the reference of campaigns, leads, and opportunities for themselves and for their cooperation teams.

i. Marketing Executive Dashboard

The dashboard is used by VPs and directors to discover if they are generating competent campaigns and generating and converting leads.

ii. Sales Executive Dashboard

Sales executives can utilize this dashboard to discover how their cooperation is doing in a particular month.

iii. Sales Manager Dashboard

Sales managers can utilize this dashboard to discover how their team is doing in a particular month.

iv. Salesperson Dashboard

Salespeople can utilize this dashboard to track their sales against the quota.

B. Customer Service CRM Dashboard.

These dashboards make it easy for people in different service and support roles to keep track of customer cases for their cooperation and agents, involving trends and key performance indicators.

i. Service Executive Overview Dashboard

This dashboard is operated by service executives to achieve prominence in their daily operations. How are the cooperation handling support issues? What are the risks? Which areas need attention? Are the goals meeting by the organizations? It also ascertains at-risk accounts, something that is imperative for executives.

ii. Agent Supervisor Overview Dashboard

It is similar to the Service Executive Overview Dashboard, but focuses on the line managers. A similar visibility at the team level is produced by this dashboard and important agent-level metrics are also ascertained.

iii. Service KPIs Dashboard

This dashboard is targeted at service executives and produces awareness of important trends, which helps them make staffing and training decisions. Executives can also observe performance against set key performance indicators (KPIs), which helps them refine cooperation goals on the basis of past performance.

IV. Research Methodology and Results

A. Research Design

In order to find out the relation between the efficiency of CRM software and the client relationship it has helped maintain, a descriptive research is conducted based on survey questioning.

B. Corpus for Data Collection

This research is based on primary data. A survey questionnaire was circulated via email in business organizations. The questionnaire is structured and based on the questions relating to the efficiency of the CRM and the client relationship. The respondents were asked to choose only one option.

C. Sample Size

Business organizations are the sampling units. The total sample size was fifty. Banks and Multinational national companies participated in this survey. The respondent score is then statistically analyzed using chi-square test in MIINITAB. The result of the analysis is the basis for the findings.

D. Research Hypothesis

A hypothesis was formulated for the CRM efficiency and the client relationship.

Null Hypothesis ($H_\circ$)	=	The Efficiency of CRM software and client relationship is independent on each other
Alternate Hypothesis (H_1)	=	The Efficiency of CRM software and client relationship is dependent on each other

E. Chi Square Test Analysis and Interpretations

Table I shows the respondents feedback to the questionnaire regarding the efficiency of CRM and the client relationship.

TABLE I
RESPONDENTS FEEDBACK OF THE QUESTIONNAIRE

Efficiency of CRM (%)	Client Relationship offered by CRM			
	Very Good	Good	Average	Not at all good
80-100	8	4	2	1
60-80	1	8	5	3
40-60	1	1	12	4

The chi Square test was conducted on Table 1 results and the analysis of chi square test is shown in Figure 8. Category 1 is 80-100%, category 2 is 60-80% and category 3 is 40-60%. It is shown that at 5% level of significance (α), the overall chi square value is 25.607 and p-value is =0.000. As the p-value is less than α, we reject the null hypothesis and take this to mean that the Efficiency of the CRM in business organizations and client relationship offered by CRM are dependent on each other and that the result is statistically significant.

Chi-Square Test: Very Good, Good, Average, Not at all good

Expected Counts are printed below observe counts

Chi-Square contributions are printed below expected counts

	Very Good	Good	Average	Not at all good	Total
1	8	4	2	1	15
	3.00	3.90	5.70	2.40	
	8.333	0.003	2.402	0.817	
2	1	8	5	3	17
	3.40	4.42	6.46	2.72	
	1.694	2.900	0.330	0.029	
3	1	1	12	4	18
	3.60	4.68	6.84	2.88	
	1.878	2.894	3.893	0.436	
Total	10	13	19	8	50

Chi-Sq = 25.607, DF = 6, P-Value = 0.000

Figure 8 Chi square results on Minitab

The chart of expected and observed values and contribution to the chi square value by category for very good, good, average and not at all good are shown in Fig 9(a), 9 (b), 9(c), 9(d) respectively. It is shown that for 'Very Good', the contribution of chi square value is highest for 80-100%. This is because of the fact that the difference between the expected and observed value is large in this category.

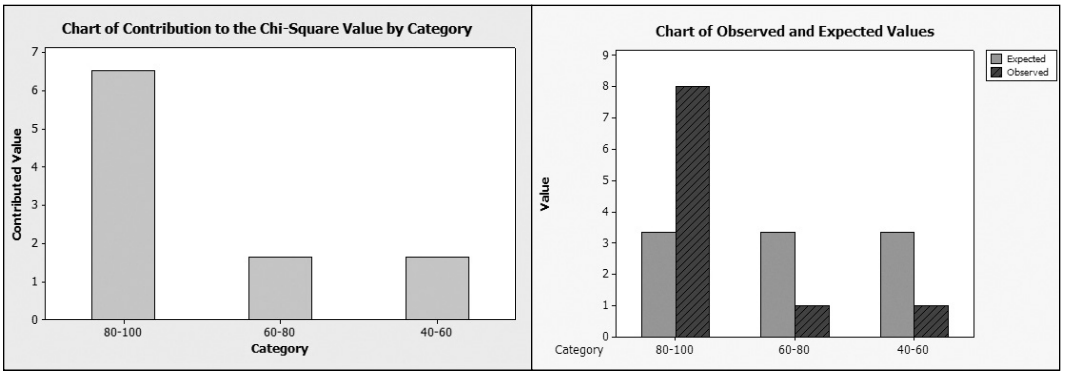


Figure 9 (a) Goodness of Fit charts for 'Very Good'

For 'Good' the highest chi square distribution is shown for the 60-80% category and the difference between expected and observed value is largest in this category as indicated in Figure 9(b).

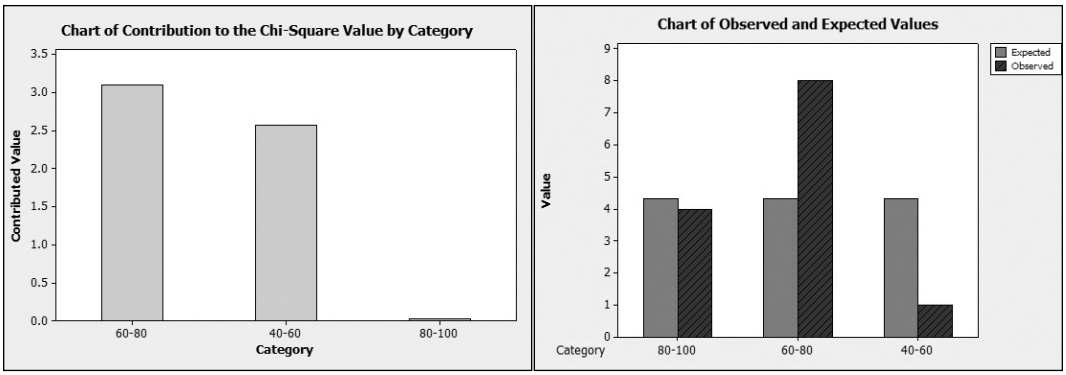


Figure 9 (b) Goodness of Fit charts for 'Good'

For the 'Average' case the highest contribution is in category 40-60% and so the difference is largest between the expected and observed value. For the 'not at all' good the chi square value is highest in the 80-100% category and the difference between the expected and observed value is largest.

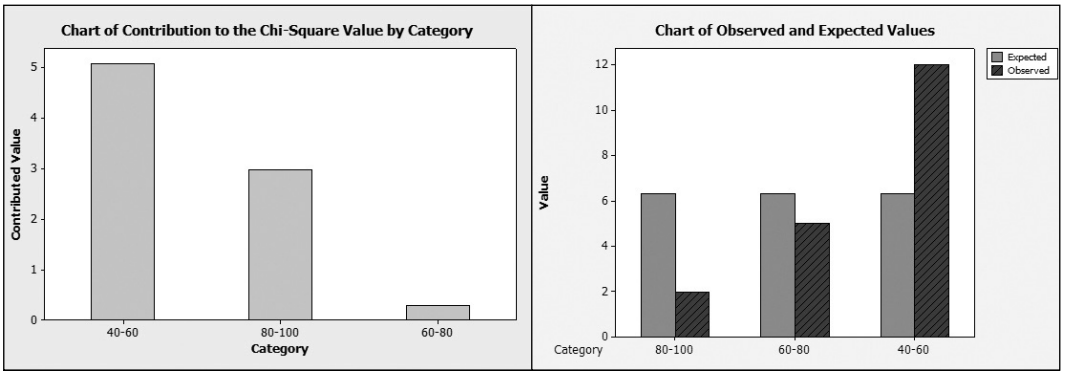


Fig 9(c) Goodness of Fit charts for 'Average'

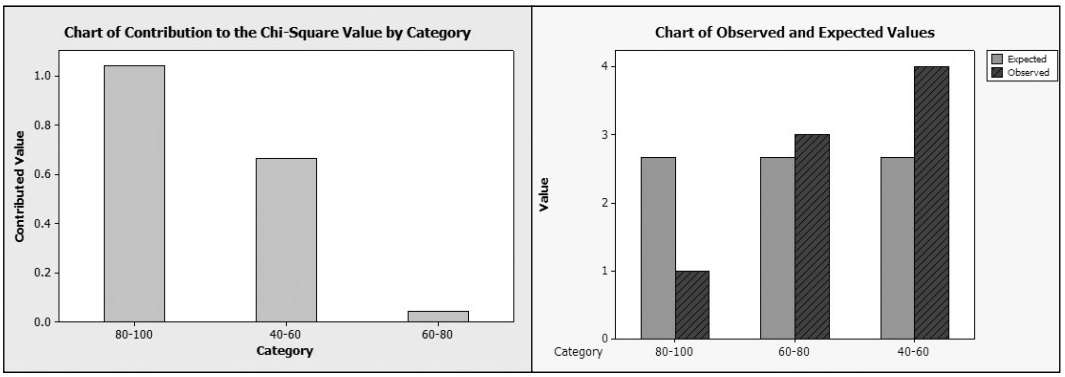


Fig 9(d) Goodness of Fit charts for 'Not at all'

V. Conclusion

Overall we believe the CRM software does help in implementing CRM processes and in managing customer relationship and loyalty by producing the complete analysis of your sales and business contacts, clients, contract wins and sales leads. The software is not utilized to its maximum as it has the potential to become more efficient. The software can be utilized in the cooperation as vital software for customer relationship management as the company's main purpose is to construct customer relationships and loyalty. Another department, where this software can be targeted, is trading desk or Sales (institutional and retail). It can be utilized in the sales department achieving the core purpose of the department i.e. execute trade in respective clients' accounts. The chi test results showed that the CRM efficiency and client relationship are dependent on each other with the Goodness of fit graphs implying that the overall contribution to the chi square value for 'very good' is highest for category '80-100'%.

References

[1] L. Ryals and S. Knox, "Cross-Functional Issues in the Implementation of Relationship Marketing Through Customer Relationship Management," European Management Journal, vol. 19, p. 534-542, 2001.

[2] A. Croteau and P. Li, "Critical success factors of CRM technological initiatives," Canadian Journal of Administrative Sciences, vol. 20, pp. 21-34, 2003.

[3] R. K. Srivastava, T. A. Shervani, and L. Fahey, "Marketing, Business Processes, and Shareholder Value: An Organizationally Embedded View of Marketing Activities and the Discipline of Marketing," Journal of Marketing, vol. 63, pp. 168-179, 1999.

[4] M. Hsieh, "A case of managing customer relationship management systems: Empirical awares and lessons learned," International Journal of Information Management, vol. 29, pp. 416-419, 2009. //

[5] Z. Lun, L.Jinlin, and W. Yingying, "Customer relationship management system framework design of Beijing Rural Commercial Bank," in IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI 2008), 2008, pp. 97-101..

[6] G. Zhang, Y. Chen, and C. Fu, "A study on the relation between business competitive benefit and CRM based on data mining," in International Technology and Innovation Conference (ITIC 2006), 2006, pp. 1710-1714.

[7] M. Maleki and D. Anand, "The Critical Success Factors in Customer Relationship Management (CRM) (ERP) Implementation," Journal of Marketing & Communication, vol. 4, pp. 67-80, 2008.

[8] F. Buttle, Customer relationship Managemen: Concepts and Tools: Oxford: Elsevier, 2004.

[9] M. Raisinghani, "CRM systems in German hospitals: illustrations of issues & trends," Journal of Cases on Information Technology, vol. 7, pp. 1-26, 2005.

[10] D. Gefen and C. M. Ridings, "Implementation Team Responsiveness and Utilizer Evaluation of Customer Relationship Management: A Quasi-Experimental Design Study of Social Exchange Theory," Journal of Management Information Systems, vol. 19, pp. 47-69, 2002.

[11] J. Dyche, The CRM Handbook: A Business Guide to Cusomer Relationship Management: Addison-Wesley, 2002.

[12] M. J. Tarokh and H. Ghahremanloo, "Intelligence CRM: A Contact Center Model," in Service Operations and Logistics, and Informatics, 2007. SOLI 2007. IEEE International Conference on, 2007, pp. 1-6.

[13] T. S. H. Teo, P. Devadoss, and S. L. Pan, "Towards a holistic view of customer relationship management (CRM) implementation: A case study of the Housing and Development Board, Singapore," Decision Support Systems, vol. 42, pp. 1613-1627, 2006.

[14] A. Keramati, N. Mojir, and H. Mehrabi, "Prioritizing Investment in CRM Resources to Improve Performance: An Empirical Investigation," in Information Management and Engineering, 2009. ICIME '09. International Conference on, 2009, pp. 330-334.

[15] A. S. Lo, L .D. Stalcup, and M. A. Lee, "Customer Relationship Management for Hotels in Hong Kong," International Journal of Contemporary Hospitality Management, vol. 22, 2009.

[16] E. J. Ragins and A. J. Greco, "Customer Relationship Management and E-Business: More Than a Software Solution," Review of Business, vol. 24, pp. 25, Winter2003 2003.

[17] M. Geib, A. Reichold, L. Kolbe, and W. Brenner, "Architecture for Customer Relationship Management Approaches in Financial Services," in the 38th Annual Hawaii International Conference on System Sciences (HICSS '05), 2005, pp. 240b-240b.

[18] K J. Trainor , J (Mick) Andzulis, A Rapp, R Agnihotri, "Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM", Journal of Business Research,Elsivier,pp.1-8, 2013

[19] Prof. Dr. Tarek Taha Ahmed, "AN EMPIRICAL EXAMINATION OF CUSTOMER RELATIONSHIP MANAGEMENT (CRM) IMPLEMENTATION IN ISLAMIC BANKS", International Journal of Academic Research and Reflection,vol.4, pp. 14-25,2016

[20] Kamrul Islam Shaon, S. M., Rahman, H., "A THEORETICAL REVIEW OF CRM EFFECTS ON CUSTOMER SATISFACTION AND LOYALTY",Parague Economic Papers, vol.4,pp 23-36, 2015

[21] Regina Reicher , Nándor Komáromi, Ágnes Szeghegyi, "The Possible Success Factors of Introduction of CRM System at Hungarian SMEs", Acta Polytechnica Hungarica,vol.12, pp.215-119,2015.

[22] [22] Rogaye Rezaeegiglo, Ali Sadouni , Feraghi Aref , Pourjedi Khotbesara, Nazari Eslam, "Review and Rating factors affecting the Deployment of (CRM) Customer Relationship Management at Nestle Company", International Journal of Academic Research in Business and Social Sciences, vol.4, pp. 539-547, 2014.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Impact of Big Data over Telecom Industry

Muhammad Waqar Khan[1], *Muhammad Asghar Khan[1], Muhammad Alam[1, 2], Wajahat Ali[1]
[1]Institute of Business and Management (IoBM), Korangi Creek, Karachi-75190, Pakistan
[2]Institute of Post Graduate Studies,Universiti Kuala Lumpur, Jalan Sultan Ismail,10250,Kuala Lumpur, Malaysia

***Abstract*- During past few years, data is growing exponentially attracting researchers to work o a popular term, the Big Data. Big Data is observed in various fields, such as information technology, telecommunication, theoretical computing, mathematics, data mining and data warehousing. Data science is frequently referred with Big Data as it uses methods to scale down the Big Data. Currently more than 3.2 billion of the world population is connected to internet out of which 46% are connected via smart phones. Over 5.5 billion people are using cell phones. As technology is rapidly shifting from ordinary cell phones towards smart phones, therefore proportion of using internet is also growing. There is a forecast that by 2020 around 7 billion people at the globe will be using internet out of which 52% will be using their smart phones to connect. In year 2050 that figure will be touching 95% of world population. Every device connect to internet generates data. As majority of the devices are using smart phones to generate this data by using applications such as Instagram, WhatsApp, Apple, Google, Google+, Twitter, Flickr etc., therefore this huge amount of data is becoming a big threat for telecom sector. This paper is giving a comparison of amount of Big Data generated by telecom industry. Based on the collected data we use forecasting tools to predict the amount of Big Data will be generated in future and also identify threats that telecom industry will be facing from that huge amount of Big Data.**

Keywords: Big Data, Data Science, Telecommunication.

I. Introduction

Big data analytics is not just a passing trend; it is becoming an important part in every aspect of a communication service [1]. Whenever we talk about the electronic communication it means we are talking about the production of data by using wired or wireless medium. The quick growth of internet and the availability of technology everywhere, internet users are increasing day by day. On the other hand, the boom in smartphone industry makes it easier for users to access network on the move. This ease of access is becoming great threat for telecom industry. People are diverting from wired to wireless medium especially on GSM to access the internet [2]. Mobile applications such as Instagram, WhatsApp, Apple, Google, Google+, Twitter, Flickr etc., providing ease to connect, at the same time are generating huge amount of data to tackle with.

*muhammad.asghar@iobm.edu.pk

Figure 1 is representing the internet traffic trends from year 2015 to 2020 [3]. Figure 2 is elaborating it more in terms of internet users against world population in percentage, growth in using internet by using smartphones, average speed of internet and average speed of per Capita per month [2].

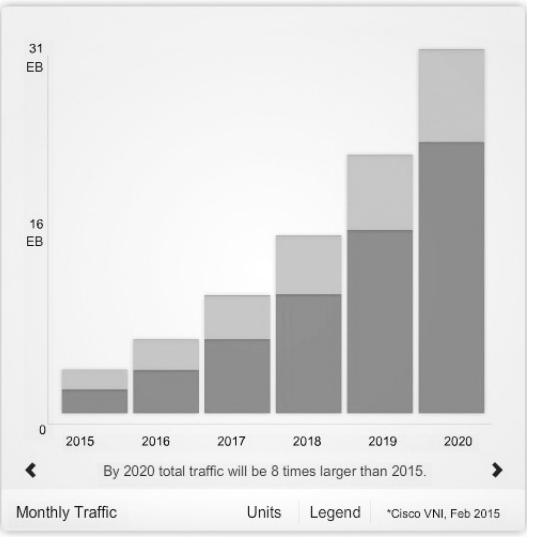


Figure 1: Internet traffic from 2015 to 2020 [3]

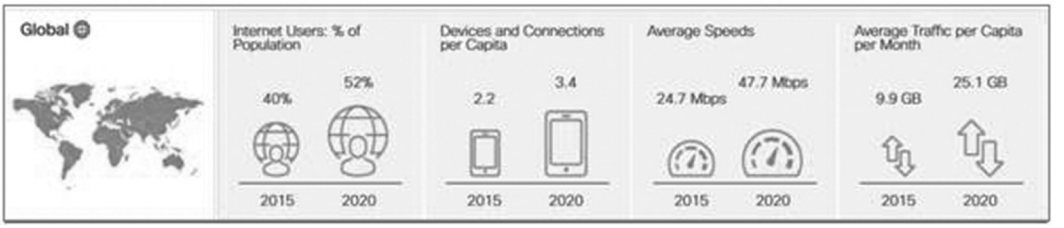


Figure 2: Internet users against world population in percentage, growth in using internet by using smartphones, average speed of internet and average speed of per Capita per month [2]

This paper is further divided into six sections; section 2 is presenting the results of previous research in Big Data and its impact over telecom industry. Section 3 is addressing the problem in the context of Big Data handling in telecom sector. Section 4 is presenting the forecasting and prediction of Big Data usage in telecom sector in next 5, 10 and 30 years. Last section concludes this paper with future directions.

II. Literature Review

This section is divided into two subsections, which are

A. Big Data

B. Big Data in Telecom Sector

A. Big Data

In [4], authors talk about the technologies, opportunities and challenges of Big Data. According to them Big Data management and analysis can be done by considering five (5) common issues which are volume, variety, velocity, value and complexity [4, 5, 6]. Volume of Big Data is normally huge. However, huge amount of data in petabytes or zettabytes is not required. In order to cope with volume problem online additional storage can be managed. Variety is another issue in Big Data management. Different type of data is produced by different devices it need to be fault tolerant. Researchers used different data mining algorithms to deal with variety. Velocity means speedup data. Normalize data provides the velocity. Authors also suggested a life cycle for Big Data by using terminologies and technologies used for Big Data. Stages they include for the life cycles are collection, filtering, analysis, storage, publication, retrieval and discovery [4].

One of the biggest problems in the Big Data is the unstructured data, which is well known as "human information". Financial records, scientific computation and simulation, videos, still images, geospatial images, weather records, call centers data generated by Facebook, tweeter, email and WhatsApp etc. all fall under this category [7]. Unstructured data makes up to 80% of all of the Big Data [7, 8]. Out of all Big Data, 84% of the data is unstructured not modeled and random. It is very difficult to analyze this huge amount of unstructured data [7].

Volume of the Big Data is becoming an alarming threat over the last past few years. Around 5 billion people over the globe call, text, tweet, send emails and browse on mobile devices [9]. In year 2012 around 89 billion emails were received and sent daily and it is forecasted that it will increase at the rate of 13% and by 2020, this amount will be touching 143 billion [10]. In year 2012, 34% of all emails were sent through smart phones [10], and this number is also increasing rapidly with the growth of and usage of smart phones [4, 11-18].

B. Big Data in Telecom Sector

Approximately 40% of the world population has internet connection today. The number of internet users has increased dramatically from 0.4 Billion in 2000 to 3.4 Billion in 2016 [19]. According to 2016 edition of Mary Meeker's annual internet trend report [20], China has the biggest internet population followed by India and USA. Total internet traffic has increased enormously in

the last two decades. In 1992, internet transferred around 100 GB of data every day, whereas these figures shoot up to 100 gigabytes per second (GBps) in 2002. Furthermore, total internet traffic reached above 20,000 GBps in 2015 [21].

Mobile computing has changed the landscape of workplace dramatically and brought many advantages to internet users such as wireless connectivity, increased productivity and easy mobility. Google Consumer Barometer [22] shows positive trends towards acceptance of mobile devices both in developed and developing countries. For example, in China and USA, the percentage of people who use a smart phone for internet has increased from 33% and 44% in 2012 to 79% and 72% in 2016 respectively. Smart phones and tablets sales have already surpassed the Workstation/PC sale. Internet live statistics [23] counter shows that more than 1.6 Billion smart phones sold in the first eight months of 2016 compared to 142 Million computers.

Applications of big data are countless and organizations of all sizes are looking to improve their performance by using big data analytics tools [24]. Yin and Kaynak [25] identified many challenges faced by enterprises when dealing with big data. One of the key challenges for the IT professionals and researcher is to deal with the application that continuously generates large quantity of data at very high speed. Enormous growth has been observed in the data production during the last two years, in fact, 90% of the world's current data has been generated in the last two years alone [26]. Massive amount of data is continuously generated and transferred through network by various internet entities, such as, video on demand, news streaming and e-commerce websites. In the era of social networking, websites are producing data at enormously high speed. For example, Twitter generates over 350,000 tweets per minute and 500 million tweets per day on average [27]. Similarly, Google processes around 40,000 search queries in every second on average [28]. Likewise, astronomy effort, such as Palomar Transient Factory in Southern California, is looking for new phenomena in the sky and roughly captures 30 terabytes of data every night [29].

Mobile devices produce huge quantity of data at enormously high speed. Cisco has developed different forecasting tools [30], based on their Visual Networking Index(VNI), that predict that "Global mobile traffic will grow three times faster than Global fixed IP traffic from 2015 to 2020". Furthermore, "Global mobile data traffic will grow 8-fold from 2015 to 2020, a compound annual growth rate of 53%". This massive shift of IP traffic from fixed to wireless network will introduce number of challenges for Network architects. Telecom companies are sitting on a gold mine, as they have plenty of data. But what they require is a proper analysis of both structured and unstructured data to get deeper insights into customer behavior, their service usage and interests real-time [31].

Since, half of the data is now being transferred via mobile and in 2012, according to study 90% of the data will be transferred through mobile and tablet devices (wireless connections). It could be a threat for telecom industry as they have already invested millions of dollars to process and store the big data via wire line connections. Now the technology is getting changed and majority of the data

is being transferred on wireless connections [32]. But just to sustain in the market, telecom industry needs to reinforce their strategies based on the current scenario. They also have an opportunity to target this untapped market. The telecom industry has an advantage over other industries due to the maximum bandwidth and depth of data it collects in the course of normal business. For example, an operator serving 8 million prepaid mobile subscribers generates around 30 million Call Data Records daily, equaling 11 billion records annually. If the same operator also provides postpaid and fixed lines services, then there is even more volume and variety of data at the ready. Big Data helps telecom operators improve their marketing effectiveness. Relevant actions can be taken based on real-time information without the need to wait for data extraction or manual data mining. Big Data can help gather real-time customer satisfaction information through social media listening or Voice of the Customer analysis. It can also improve customer experience by identifying the most valuable customers who would benefit from dedicated treatment and better services [32].

Different subscription fees can be charged to customer according to their usage on monthly basis. Cloud based services could help telecommunication industries to manage and secure customer's data. Telecom Industry can create new products just to utilize the big data with wireless connections. Top players like Google, Skype, and Netflix may be interested in behavioral data to target consumers with specific content.

III. Problem Statement

Technology is rapidly shifting from wired to wireless networks and every node is busy to generate data. This technology shift in data growth is predicting that we will face serious infrastructural challenges, which include but not limited to, bandwidth problem, network maintenance and support issues, computational bottleneck and MSC's and BSC's congestions. Currently organizations that are dealing with Big Data are increasing resources for load balancing. However, increasing the number of IT resources might solve the problem for the short run but this approach is not scalable for the future growth. IT companies, such as Google, Facebook and BT are looking for some improved protocols and algorithms to improve data transmission and computing. Forecasting the future data growth/transmission in short and long run will help IT researchers, Telecom companies, Governments and other stakeholders to understand where they currently stand and where and in which direction they need to move to solve above mentioned problem. This research paper is using Time Series forecasting to understand the data needs in 2020, 2030, 2040 and 2050.

IV. Proposed Methodology

We have collected data for past 16 years from year 2000 to 2015 of world population and compare it with internet and non-internet users. Based on the available data by using regression we predict for 2020, 2030, 2040 and year 2050 as shown in table 1. Rapid growth of internet users is showing that the amount of data will be immense in year 2050. Trends shows that most of the data will be generated by cellphones/smartphones, because of their handy and available every time nature. This will overburden telecom industry because it will be dealing a huge amount of Big Data. Current research shows that more than 52% of users are connected via smartphones as compare to 41% via desktop [36].

Table 1
Year wise distribution of world population, internet and non-internet users

Year	Internet Users	World Population	Non Users
2000	414,794,957	6,126,622,121	5,711,827,164
2001	502,292,245	6,204,310,739	5,702,018,494
2002	665,065,014	6,282,301,767	5,617,236,753
2003	781,435,983	6,360,764,684	5,579,328,701
2004	913,327,771	6,439,842,408	5,526,514,637
2005	1,030,101,289	6,519,635,850	5,489,534,561
2006	1,162,916,818	6,600,220,247	5,437,303,429
2007	1,373,226,988	6,681,607,320	5,308,380,332
2008	1,575,067,520	6,763,732,879	5,188,665,359
2009	1,766,403,814	6,846,479,521	5,080,075,707
2010	2,023,202,974	6,929,725,043	4,906,522,069
2011	2,231,957,359	7,013,427,052	4,781,469,693
2012	2,494,736,248	7,097,500,453	4,602,764,205
2013	2,728,428,107	7,181,715,139	4,453,287,032
2014	2,956,385,569	7,265,785,946	4,309,400,377
2015	3,185,996,155	7,349,472,099	4,163,475,944
2016	**3,424,971,237**	**7,432,663,275**	**4,007,692,038**
2017	3,448,735,286	7,507,238,701	4,058,503,415
2018	3,640,880,317	7,589,116,161	3,948,235,844
2019	3,833,025,348	7,670,993,621	3,837,968,273
2020	**4,025,170,380**	**7,752,871,081**	**3,727,700,702**
2021	4,217,315,411	7,834,748,541	3,617,433,130

2022	4,409,460,443	7,916,626,001	3,507,165,558
2023	4,601,605,474	7,998,503,461	3,396,897,987
2024	4,793,750,505	8,080,380,921	3,286,630,416
2025	4,985,895,537	8,162,258,381	3,176,362,844
2026	5,178,040,568	8,244,135,841	3,066,095,273
2027	5,370,185,600	8,326,013,301	2,955,827,701
2028	5,562,330,631	8,407,890,761	2,845,560,130
2029	5,754,475,662	8,489,768,221	2,735,292,559
2030	**5,946,620,694**	**8,571,645,681**	**2,625,024,987**
2031	6,138,765,725	8,653,523,141	2,514,757,416
2032	6,330,910,757	8,735,400,601	2,404,489,844
2033	6,523,055,788	8,817,278,061	2,294,222,273
2034	6,715,200,820	8,899,155,521	2,183,954,701
2035	6,907,345,851	8,981,032,981	2,073,687,130
2036	7,099,490,882	9,062,910,441	1,963,419,559
2037	7,291,635,914	9,144,787,901	1,853,151,987
2038	7,483,780,945	9,226,665,361	1,742,884,416
2039	7,675,925,977	9,308,542,821	1,632,616,844
2040	**7,868,071,008**	**9,390,420,281**	**1,522,349,273**
2041	8,060,216,039	9,472,297,741	1,412,081,702
2042	8,252,361,071	9,554,175,201	1,301,814,130
2043	8,444,506,102	9,636,052,661	1,191,546,559
2044	8,636,651,134	9,717,930,121	1,081,278,987
2045	8,828,796,165	9,799,807,581	971,011,416
2046	9,020,941,196	9,881,685,041	860,743,845
2047	9,213,086,228	9,963,562,501	750,476,273
2048	9,405,231,259	10,045,439,961	640,208,702
2049	9,597,376,291	10,127,317,421	529,941,130
2050	**9,789,521,322**	**10,209,194,881**	**419,673,559**

A. Regression Analysis of time series analysis using SPSS:

Linear trend is found between time and Internet users. Time is considered as independent variable (x) and Internet users as dependent variable. The equation describe as [33-35]:

Y (internet users) = - 3.841 x 1011 +1.92x108 x (time)

According to regression line estimated internet user would be raised 5.95 billion in 2030 and 9.79 billion in 2050; this is 69.4% and 95.5% of the total world population respectively. Following graph describe Linear regression support trend line and highly significant. The correlation r=0.992 define highly positive relationship between time and internet users. Figure 3 and 4 are representing trends of internet users till year 2050. Figure 5 is presenting the user inclination towards cellphones. 2015 data shows that 51% of users are using smartphones as compare to desktop users which were only 42% and dropping [36].

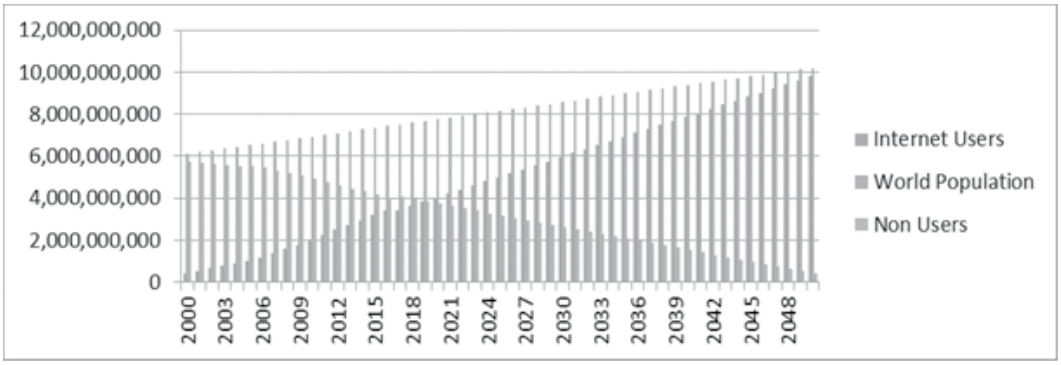


Figure 3: Bar chart representation of world population, internet users and non-internet against time.

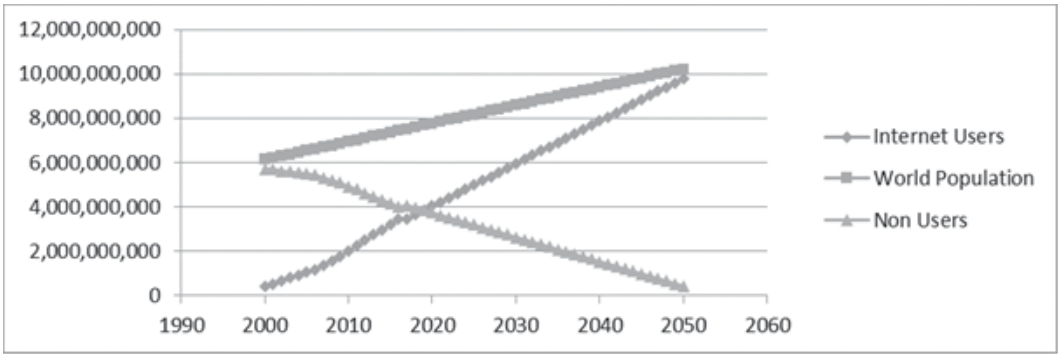


Figure 4: Trend line of world population, internet users and non-internet against time.

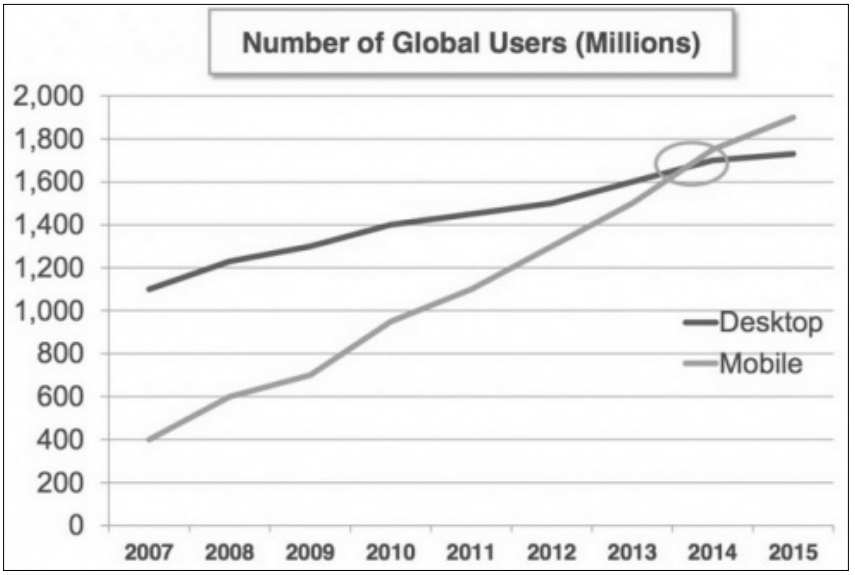


Figure 5: Figure by Dave Chaffey, showing the increase in use of mobile phones over desktop over the years [36]

V. Conclusion, Limitation and Future Directions

Forecasting by using regression shows that in year 2050 95% of world population will be connected via internet and most of it will be using wireless mode to connect especially GSM or future telecom architecture. It will generate huge amount of data for telecom industry for which the industry is not ready to deal with. Most of this data will be redundant therefore proper data mining tools and techniques are required to dig out for the required data and dump the redundant data.

References

[1] Big data analytic is just not a passing trend. http://www.analysysmason.com/Research/Content/Reports/Big-data-analytics-Jul2014/ (last visited: September 16th, 2016)

[2] http://www.cisco.com/c/m/en_us/solutions/service-provider/vni-forecast-highlights.html (last visited: September 16th, 2016)

[3] http://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-widget/forecast-widget/advanced.html (last visited: September 16th, 2016)

[4] Khan, N., Yaqoob, I., Hashem, I.A.T., Inayat, Z., Mahmoud Ali, W.K., Alam, M., Shiraz, M. and Gani, A., 2014. Big data: survey, technologies, opportunities, and challenges. The Scientific World Journal, 2014.

[5] S.Kaisler, F.Armour, J.A.Espinosa, W.Money, "Big Data: issues and challenges moving forward", Proceedings of the IEEE 46th Annual Hawaii International Conference on System Science (HICSS'13), pp.995-1004, January 2013.

[6] S.Sagiroglu, D.Sinanc, " Big Data: A review" Proceedings of the international onference on Collaboration Technologies and Systems (CTS'13), pp 42-47, IEEE, San Diego, Calif, USA. May 2013.

[7] K.Douglas, "Infographic: Big Data brings marketing baig numbers" 2002, http://www.marketingtechblog.com/ibm-

big-data-marketing/. (Last visited: August 20, 2016)

[8] A.Holzinger, c.Stocker, B.Ofner, G.Prohaska, A.Brabenetz and R.Hofmann-Wellenhof, "Combining HCI, natural language processing and knowledge discovery- potential of IBM content analytics as an assistive technology in the biomedical field", in Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured Big Data, vol. 7947 of Lecture Notes in Computer Science, pp. 13-24, Springer, Berlin, Germany, 2013.

[9] P.Russom, "Big Data analytics", TDWI Best Practices Report, Fourth Quarter 2011.

[10] S.Radicati and Q.Hoang, " Email Statistics Report", 2012-2016. The Radicati Group, London, UK, 2012.

[11] Worldmeters, "Real time world statistics", 2014 http://www.worldmeters.info/world-population/. (Last visited: August 22, 2016).

[12] M.Chen, S.Mao and Y.Liu, "Big Data: A Survey", Mobile Networks and Applications, Vol. 19, No. 2, pp. 177-209, 2014.

[13] S.Kaisler, F.Armour, J.A.Espinosa and W.Money, "Big Data: Issues and Challenges moving forward", in Proceedings of the IEEE 46th Annual Hawaii international Conference on System Sciences (HICSS'13), pp.995-1004, January 2013.

[14] A Muhammad, "Power Management of Portable Devices by Using Clutter Based Information" IJCSNS, International Journal of Computer Science and Network Security, VOL. 9 No. 4, April 2009, tel-00815919, version 1-19 Apr 2013 pp 237-244.

[15] A Muhammad, MS Mazliham, M Shahrulniza, M Amir, "Posterior Probabilities based Location Estimation (P2LE) Algorithm for Locating a Mobile Node in a Disaster Area M" MULTICONF p 13-16, 2009.

[16] Muhammad Alam, Mazliham Muhammad Suud, Patrice Boursier, Shahrulniza Musa, Jawahir Che Mustapha Yusuf, "Predicted and corrected location estimation of mobile nodes based on the combination of kalman filter and the bayesian decision theory" , International Conference on Mobile Wireless Middleware, Operating Systems, and Applications, Pages 313-325, Publisher Springer Berlin Heidelberg

[17] A Muhammad, MS Mazliham, Patrice Boursier, M Shahrulniza, Jawahir Che Mustapha Yusuf, "Terrain/clutter based location prediction by using multi-condition Bayesian decision theory" Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication. P. 110-117, 2012.

[18] Muhammad Alam, "Error corrected location determination in an outdoor wireless environment by using estimation, filtering, prediction and fusion techniques: A Wi-Fi application by using terrain based knowledge" , HAL, 2013 pages 313.

[19] http://www.internetlivestats.com (last visited: October 19th, 2016)

[20] http://www.kpcb.com/blog/2016-internet-trends-report (last visited: October 19th, 2016)

[21] The Zettabyte Era—Trends and Analysis, Cisco, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html (last visited: September 27th, 2016)

[22] https://www.consumerbarometer.com/en/trending/?countryCode=CN&category=TRN-NOFILTER-ALL (last visited: September 27th, 2016)

[23] http://www.internetlivestats.com (last visited: September 27th, 2016)

[24] BUILDING A SCALABLEBIG DATA INFRASTRUCTURE FOR DYNAMIC WORKFLOWS,

[25] http://www.emc.com/collateral/solution-overview/h11761-building-scalable-big-data-infra-so.pdf (last visited: September 27th, 2016)

[26] YIN, BS, Kaynak, O (2015),Big Data for Modern Industry: Challenges and Trends, Proceedings of the IEEE, vol. 103, No. 2.

[27] Mukerjee, S and Shaw, R 2016, Big Data – Concepts, Applications, Challenges and Future Scope, International Journal of Advanced Research in Computer and Communication Engineering, vol. 5, no. 2.

[28] http://www.internetlivestats.com/twitter-statistics/ (last visited: September 27th, 2016)

[29] http://www.internetlivestats.com/google-search-statistics/ (last visited: September 27th, 2016)

[30] http://www.nytimes.com/2015/08/01/science/research-scientists-to-use-network-much-faster-than-internet.html (last visited: September 27th, 2016)

[31] VNI Complete Forecast Highlights Tool (2016), http://www.cisco.com/c/m/en_us/solutions/service-provider/vni-

forecast-highlights.html (last visited: September 27th, 2016)

[32] Telecom companies are sitting on a gold mine. http://bigdata-madesimple.com/11-interesting-big-data-case-studies-in-telecom/ (last visited: September 27th, 2016)

[33] Big Data helps telecom operators improve their marketing effectiveness. http://www.teletech.com/resources/articles/numbers-three-big-data-opportunities-telecoms#.V82-gvmqqko (last visited: September 27th, 2016)

[34] Rob J Hyndman and George Athana¬sopou¬los Journal of Forecasting J. Fort. 20, 315-328(2001) DOI: 10.1002/for.795

[35] Joseph Nguyen by Regression Basics for Business Analysis

[36] Statistics for Management 7th Edition by Richard Levin

[37] Dave Chaffey, "Statistics on consumer mobile usage and adoption to inform your mobile marketing strategy mobile site design and app development", Mobile Marketing Statistics compilation: Mobile Market Analytics, April 27, 2016.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Comparative Analysis of E–Shape and H-Shape Microstrip Slotted Antenna on the Basis of Return Loss and Bandwidth

[1, 2*]J. Iqbal, U. Illahi, [1]M.I. Sulaiman, [3, 4]M. Alam, [5]MS.Mazliham
[1]Universiti Kuala Lumpur, British Malaysian Institute (BMI), batu 8, Jalan Sungai Pusu, 53100 Gombak, Malaysia
[2]Gomal University, Dera Ismail khan KPK, Pakistan
[3]Institute of Business and Management (IoBM), Korangi Creek, 74900 Karachi, Pakistan
[4]Universiti Kuala Lumpur, IPS, UniKL1016, Jalan Sultan Ismail, 50250 Kuala Lumpur, Malaysia
[5]Malaysia France Institute (MFI) Universiti Kuala Lumpur, Bander Baru, 43650 Bangi, Malaysia

Abstract **-Micro-strip patch antenna is more popular as compared to other antennas and widely used today in modern wireless communication system because of its elegant weight, low fabrication price, easy to feed, ease to fabricate, easy to analyze, and attractive radiation characteristics. Though patch antenna has several advantages but has some disadvantages as well, like limited bandwidth and low gain. In last few years, Different techniques such as Electromagnetic Band Gap (EBG) structure, artificial magnetic conductor (AMC)/ and modifying the structure of microstrip patch antenna are used to overcome the above mentioned limitations. In this paper, we propose two different geometry shape antennas i.e. E and H-shaped which is developed from the rectangular patch and compare the performance of the proposed antennas on the basis of gain and Bandwidth with the same physical dimensions. The CST Micro studio simulator is used to study the antennas. Bandwidth and Return loss of E-shaped Antenna is much better than that of H-shaped antenna. The results obtained clearly demonstrations that, the implementation of the E-Shape is much better in term of bandwidth and return loss compared to the H-shaped patch antenna. So it's better to use E-shaped antenna for modern wireless communication system.**

Keywords: Microstrip patch antenna, E-Shape, H-shape, Bandwidth, Wireless communication system.

I. Introduction

Microstrip antenna is a significant Element in mostly every wireless communication and global positing system (GPS) since it was first explaining by Heinrich Hertz in 1886 and in 1901 Gulielmo Marconi demonstrate its practical application [1]. Further scientists are more interested in compact design of antenna for communication systems. In last few years, microstrip antennas have caught the attention of researchers for their application in mobile communication, satellite communication,

* javed.iqbal@s.unikl.edu.my

and wireless communication systems because of their compact elegant weight, size, and easygoing fabrication characteristics.

Microstrip antenna is utmost shared option used to understand millimeter wave monolithic integrated circuits for radar, microwave and communication purposes [2]. The fundamental Alignment of a microstrip antenna is a metallic piece printed on thin, grounded dielectric substrate [1]. Initially, the element was fed with a coaxial line over the base of the substrate, or by a coplanar microstrip line, let feed network and other circuitry to be fabricated in the same configuration as the antenna element. The microstrip antenna emitted microwaves in a comparatively broad beam broadside of to the plane of the substrate, therefore the microstrip antenna has a very low profile, and can be fabricated using printed circuit procedure [3].

The frequency of the patch antenna at which it will operate is determined by the length L of patch. The center frequency will be approximately given by According to the above equation the patch antenna should have a length equal to one half (1/2) of a wavelength within the dielectric (substrate) medium [4].

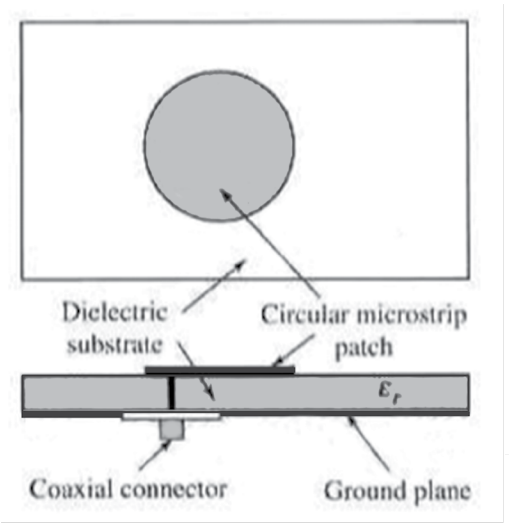
$$f_c \approx \frac{c}{2L\sqrt{\varepsilon_r}} = \frac{1}{2L\sqrt{\varepsilon_0\ \varepsilon_r\ \mu_0}} \quad (1)$$

II. Literature Survey

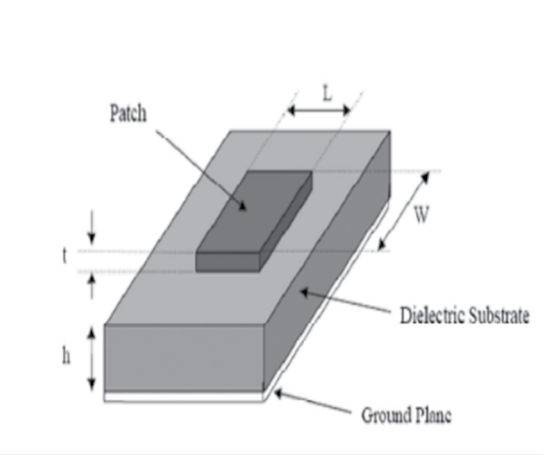
Basically, Microstrip patch antenna comprises two type of radiating patches i.e. Rectangular or circular, these patches are made up of substrate on one side and ground plane on the other side [1]. Micro strip Patch antennas have a significant role in the field of wireless communication. Usually microstrip fabrication technique is used for the construction of a microstrip patch antenna as this method is the simplest one.

The patch is made of mostly of copper which work as conducting material and can modify to any shape as shown in Figure 1 (a, b). The microstrip patch is usually triangular, circular, elliptical, square and rectangular or in some other shapes like diamond as well [4-7].

The rectangular Microstrip patch antenna is generally used for all the kinds of Microstrip antennas. It's easy to fabricate, tough design is easy to handle. When source signal is applied to Microstrip patch antenna at patch, the electromagnetic waves will be emitted. This patch is a strip of length L and width W on a dielectric substrate with constant εr, height of the patch is denoted by h and thickness is represented by t is sustained by a ground plane.



a) Circular Micro Strip Antenna



b) Rectangular Micro Strip Antenna

Figure 1. Different geometry of micro strip antenna

A. Methods to analyze performance of microstrip antenna

There are few different methods to analyze the performance of microstrip antenna the most popular parameters are:

1) Return loss

2) Bandwidth

3) Far field radiation (Directivity and Gain)

1. RETURN LOSS:

First parameter is return loss. Basically, it is a logarithmic ratio measured in dB which matches the power reflected by the antenna to the power that is fed into the antenna through the transmission line. Voltage standing wave ratio (VSWR) is a way to measure transmission line imperfections. Return loss is calculated at -10dB [1] as shown in Figure 2.

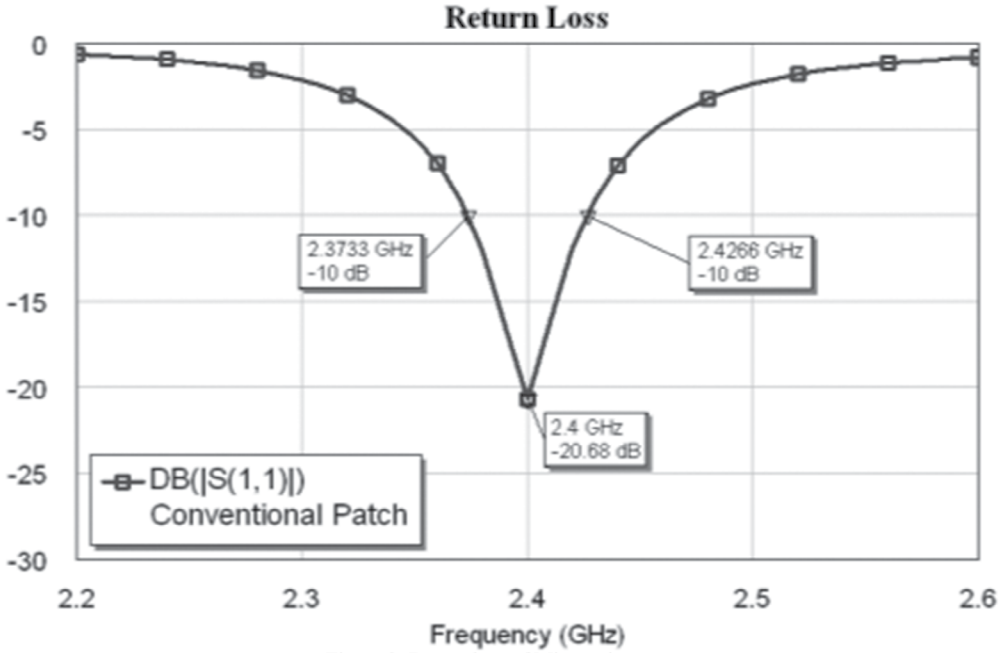


Figure 2: Return loss of Microstrip antenna

2. BANDWIDTH:

One of the limitations of the microstrip antenna is its narrow bandwidth. The bandwidth can be defined in terms of its VSWR, for circularly polarized antenna, bandwidth is expressed in terms of Axial Ratio. Bandwidth is defining inform of percentage [3].

3. FAR FIELD RADIATION (DIRECTIVITY AND GAIN):

Far field radiation parameters are mostly Gain and directivity

3.1. Gain

Antenna gain is a degree of how much more power an antenna under test (AUT) will emit in a specific direction regarding that of a dipole antenna or isotropic Emitter (dBi) [1].

3.2. Directivity

Directivity is an amount of how 'Specific' an antenna's radiation pattern is. Directivity is the gain measured supposing an isotropic antenna. Actual antennas have losses, and gain is just the directivity multiplied by the proficiency of the antenna [1].

III. Problem Statement

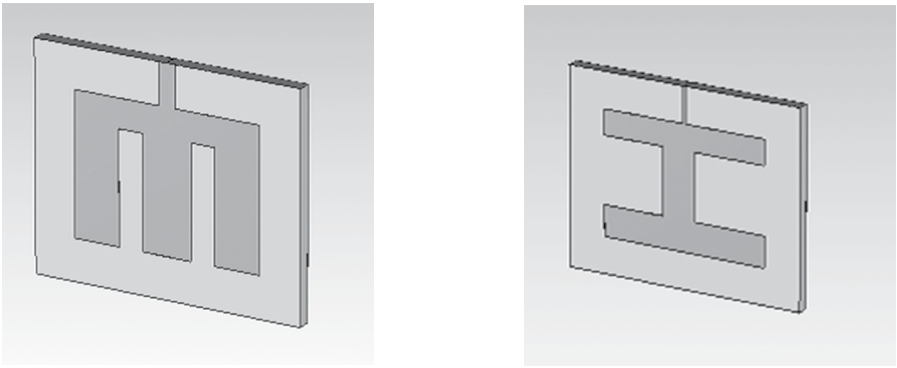
In general, Microstrip antennas resonate effectively at a specific frequency related to their major mode and having normally narrow bandwidth (1-2%) and low Directivity as well. Because of these drawbacks in conventional form of microstrip antennas fail to find much Attention in modern wireless communication [3].

Wireless communication system which must be capable of operating at two different or more frequencies at a time and must present broadband performance [6], considering this requirement, conventionally printed circuit antenna fail to serve their purpose in wireless communications system and therefore modifications in conventional microstrip antenna geometry is recommended [1] like Compact Arrays antenna [8] A Microstrip Slotted Patch Antenna Using Artificial magnetic conductor (AMC) [9-14].

Because of its low profile structure, Microstrip patch Antennas have a number of advantages However, narrow bandwidth has been the main disadvantage for this antennas. There is numerous technique which has been applied by different researcher to overcome narrow bandwidth issues, i.e. indulging parasitic elements, rising the substrate thickness, stacking configuration, introducing ring in patches or modifying the patch shape includes designing an E shaped-shaped and H-shaped antenna [1,3,5]. It has been discovered that H-slot microstrip antenna offers bandwidth up to 30% while E shaped patch antenna can raise bandwidth above 30%. In this paper, we will compare the performance of E & H Shaped slotted antenna structure on the basis of above mention parameters.

IV. Proposed Model and Design

In this design, we will study two different geometry shapes H and E slotted antenna as shown in Fig 3(a, b). The two proposed antennas are simulated using the CST STUDIO SUITE, and results matched with each other, and the shape with improved bandwidth will be adopted and implemented.



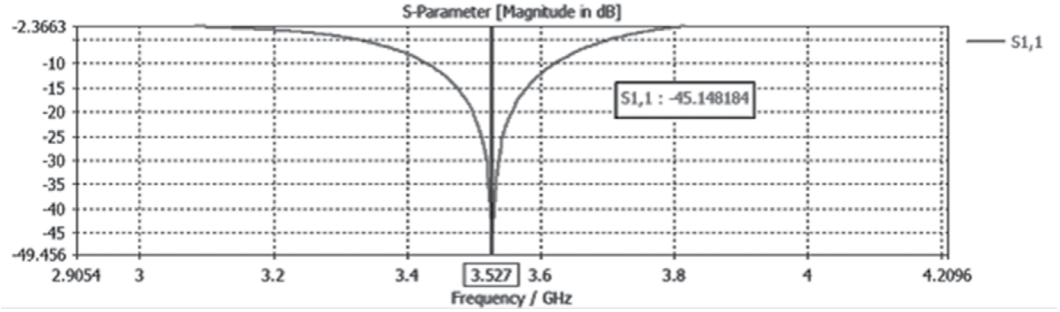
a) E-shaped slotted antenna b) H-shaped slotted antenna

Figure 3: E and H shaped microstrip antenna

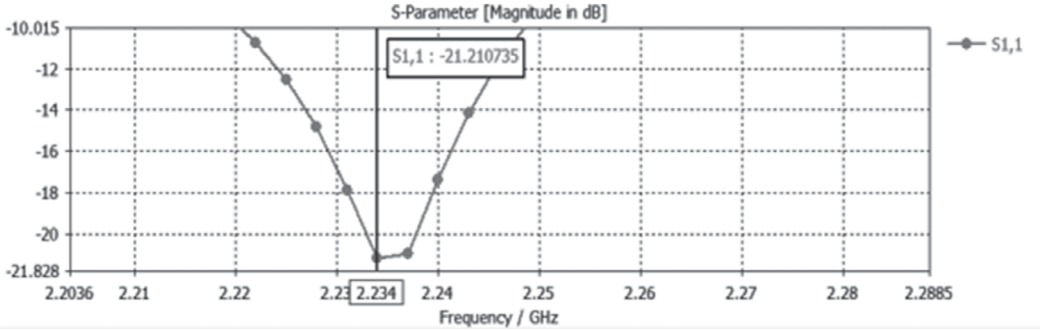
V. Results and Discussion

Figure 3 shows the geometry of E and H shaped microstrip antenna, in Fig 4(a, b) it is shown that E and H designed slotted micro strips are resonated at frequency of 3.527Ghz and 2.234Ghz Respectively.

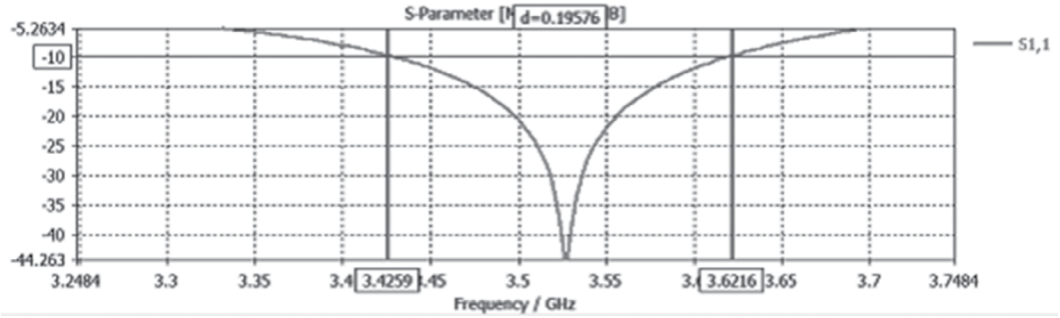
Figure 4 (a, b) shows the return loss of E shaped antenna which is -45.14dB while on the other side H- shape antenna has -21.21dB return loss, if we check out the most important parameter, bandwidth of both antenna, it is clear from the table1 that E shape antenna is far better than H- shape patch antenna, which is 5.55% and 1.27% respectively, Figure 4 (c, d).



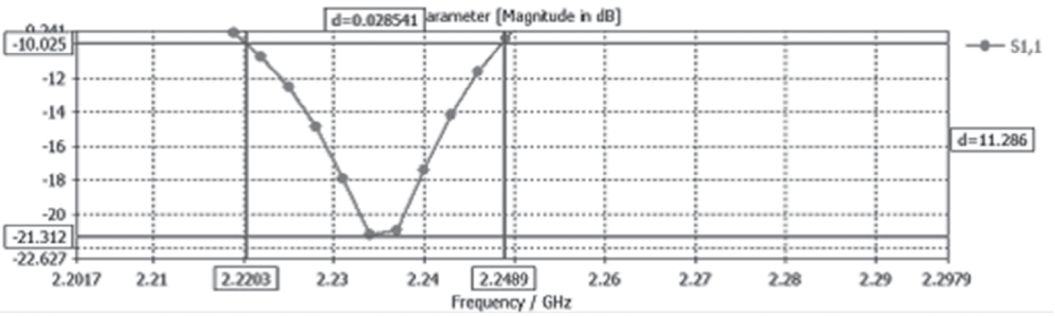
a) Return loss E shaped antenna



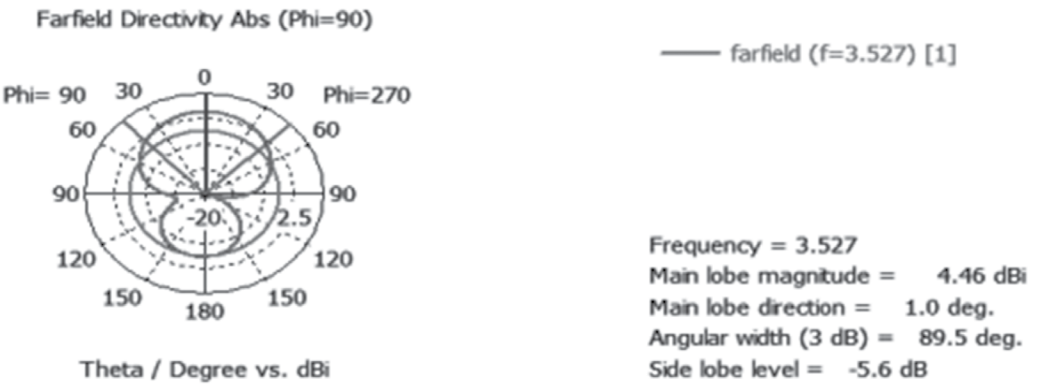
b) Return loss H shaped antenna



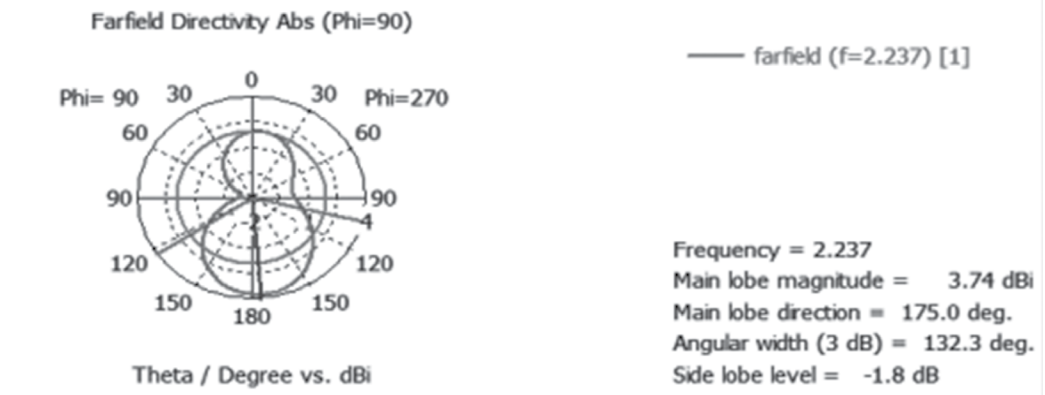
c) Bandwidth of E-shaped antenna



d) Bandwidth of E-shaped antenna



e) Directivity of E-shape antenna



f) Directivity of E-shape antenna

Figure 4 Analysis of different parameters of E and H Shape Antenna

TABLE 1
COMPRESSION ON DATA OF E AND H SHAPED ANTENNA

S/N	Type	Return loss	Bandwidth	Directivity
1	E-shape	-45.14dB	5.55%	4.46dBi
2	H-shape	-21.21 dB	1.27%	3.71dBi

VI. CONCLUSION

In this paper, the aim was to judge the performance of two antennas having different geometry. We have selected two different patch antennas and the simulated results compare with each other. The results obtained clearly shows in table 1 that bandwidth of E-shaped slotted antenna has higher than that of H-shaped patch antenna. We have observed that E-shaped patch antenna has highest Return loss as well. The reported results indicate that the proposed antenna geometry (E-shaped) Full fill all the requirement for an antenna used for wireless communication systems.

ACKNOWLEDGMENTS

The author would like to acknowledge the Research and Innovation Department of British Malaysian Institute for providing peace full environment for research and Gomal University, Pakistan for granting oversea study leave for PhD.

REFERENCES

[1] Behera, S. K. "Novel Tuned Rectangular Patch Antenna as a load for phase power combining." PhD diss., 2007.

[2] Kaur, Jaswinder, and Rajesh Khanna. "Co-axial fed rectangular microstrip patch antenna for 5.2 GHz WLAN application." Universal Journal of Electrical and Electronic Engineering 1, no. 3 (2013): 94-98. H. Poor, An Introduction to Signal Detection and Estimation. New York: Springer-Verlag, 1985, ch. 4.

[3] B. Smith, "An approach to graphs of linear forms (Unpublished work style)," unpublished.

[4] Sandeep Kumar, Suresh Sahni, Ugra Mohan Kumar, Devendra Singh," Design of Circularly Polarized Microstrip Square patch Antenna for Improved Bandwidth And Directive Gain With Low Return Loss" IJECET Volume 4, Issue 2, March – April, 2013, pp. 324-331

[5] Gupta, Sushila. "Design a H shaped Patch Antenna for Wireless Communication." International Journal of Advanced Electrical and Electronics Engineering (IJAEEE), ISSN (2013): 2278-8948..

[6] Ray, K. P. "Broadband, dual-frequency and compact microstrip Antennas." Ph. Dthesis, Indian Institute of Technology, Bombay, India (1999).

[7] Rafi, Gh Z., and L. Shafai. "Wideband V-slotted diamond-shaped microstrip patch antenna." Electronics Letters 40, no. 19 (2004): 1166-1167.

[8] Oispuu, M., and U. Nickel. "Array-based direction finding and localization for small aerial platforms." In Antennas and Propagation in Wireless Communications (APWC), 2014 IEEE-APS Topical Conference on, pp. 241-244. IEEE, 2014.

[9] Manju Saini, Sachin Singla, Koneesh Aggarwal "A Microstrip Slotted Patch Antenna Using Amc" Manju Saini et al Int. Journal of Engineering Research and Applications ISSN: 2248-9622, Vol. 4, Issue 4(Version 8), April 2014, pp.138-142.

[10] A Muhammad, "Power Management of Portable Devices by Using Clutter Based Information" IJCSNS, International Journal of Computer Science and Network Security, VOL. 9 No. 4, April 2009, tel-00815919, version 1-19 Apr 2013 pp 237-244.

[11] A Muhammad, MS Mazliham, M Shahrulniza, M Amir, "Posterior Probabilities based Location Estimation (P2LE) Algorithm for Locating a Mobile Node in a Disaster Area M" MULTICONF p 13-16, 2009.

[12] Muhammad Alam, Mazliham Muhammad Suud, Patrice Boursier, Shahrulniza Musa, Jawahir Che Mustapha Yusuf, "Predicted and corrected location estimation of mobile nodes based on the combination of kalman filter and the bayesian decision theory" , International Conference on Mobile Wireless Middleware, Operating Systems, and Applications, Pages 313-325, Publisher Springer Berlin Heidelberg

[13] A Muhammad, MS Mazliham, Patrice Boursier, M Shahrulniza, Jawahir Che Mustapha Yusuf, "Terrain/clutter based location prediction by using multi-condition Bayesian decision theory" Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication. P. 110-117, 2012.

[14] Muhammad Alam, "Error corrected location determination in an outdoor wireless environment by using estimation, filtering, prediction and fusion techniques: A Wi-Fi application by using terrain based knowledge" , HAL, 2013 pages 313.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Finite Difference Method with Dirichlet Problems of 2D Laplace's Equation in Elliptic Domain

[1*]Ubaidullah and [2]Muhammad Saleem Chandio
[1]Department of Mathematics, Sukkur Institute of Business Administration
[2]Institute of Mathematics and Computer Science, University of Sindh, Jamshoro

***Abstract*- In this study finite difference method (FDM) is used with Dirichlet boundary conditions on rectangular domain to solve the 2D Laplace equation. The chosen body is elliptical, which is discretized into square grids. The finite difference method is applied for numerical differentiation of the observed example of rectangular domain with Dirichlet boundary conditions. The obtained numerical results are compared with analytical solution. The obtained results show the efficiency of the FDM and settled with the obtained exact solution. The study objective is to check the accuracy of FDM for the numerical solutions of elliptical bodies of 2D Laplace equations. The study contributes to find the heat (temperature) distribution inside a regular rectangular elliptical discretized body.**

Keywords: Dirichlet Boundary Conditions, Finite Difference Method, Laplace Equation, Elliptic Domain.

I. Introduction

The Laplace partial differential equation in two independent variables have important applications in engineering and science, like fluid flow, electricity and steady heat conduction. In engineering and science mostly deals with variables like and to discuss space with timealso as independent variable for a modeled physical problem is considering as dependent variable. Engineers and scientist investigate the actual partial differential equations (PDE's) that given the investigated physical problem. Many numerical methods are invented in 20th century to solve Elliptic partial differential equations (EPDE's). Physical problems like sound, heat, electrodynamics, fluid flow, elasticity etc. are formulated mathematically by Partial differential equations (PDE's). The Neumann and Dirichlet boundary conditions are mostly applied to obtain the solution of 2D Laplace equation. M.L. Dhumal and S.B. Kiwne [1] used Neumann and Dirichlet boundary conditions to obtain the solution of Laplace equation. The approximate solution of two dimensional Laplace equation using Dirichlet conditions is also discussed by Parag V. Patil and J.S.V.R. Krishna Prasad [2].

Laplace equation is used to solve Cauchy problem by Qian et al [3]. The solution of Laplace equation with simple boundary conditions studied by Morales et al [4]. Lesnic *et al* [5] did work for the solution of Cauchy problem to the Laplace equation using an iterative boundary element method.

* ubaidullah@iba-suk.edu.pk

Li *et al* [6] have studied Laplace's equation on elliptic domains by using Dirichlet conditions. Laplace equation in circular domains with circular hole by using Neumann problems of Laplace's equation in circular domains with circular holes have studied by Lee *et al* [7]. Lee *et al* concluded that the method of field equation's is an effective method to solve the Neumann problems. Smith G.D. [8], Ames W.F. [9], Lapidus Land Pinder G.F. [10] and Greenspan D. and Parter S.V. [11] studied the Finite difference methods (FDM) for partial differential equations.

Many numerical methods are invented in 20th century to solve Elliptic partial differential equations. Since 1900 the applications of FDM for PDE's have been known. To solve the elliptic interface problems finite difference method is an accurate method studied by J. Thomas [12]. In 1960 the mesh based methods finite difference method and finite element method was used for numerical solutions of ODE's and PDE's. Jensen [13] worked with fully arbitrary meshes by using FDM. FDM's and FEM's are more suitable for regular meshes. Perrone and Kaos [14] worked on irregular meshes by using two dimensional FDM. P.G. Martinsson [15] discussed for variable coefficient elliptic partial differential equations discretized by composite spectral collocation method. For solving irregular domains by FEM is a relatively time consuming. Ames [16] worked to solve PDE's in irregular domains by FDM. On irregular 2D domain Orovio et al [17] studied the spectral method to solve reaction-diffusion equation. This paper used finite difference method to get the discrete numerical approximations for the derivative. Finite difference method is used here to discretize the domain into uniform grids.

In engineering elliptic partial differential equations used to describe steady-state boundary value problems. For the approximate solution of elliptic partial differential equations (EPDE's), the given partial differential equation is converted into an algebraic difference equation. In this paper we used finite difference method to determine potential in rectangular domain using Dirichlet boundary conditions.

II. Problem Formulation

$$\nabla^2 T^* = \frac{\partial^2 T^*}{\partial x^2} + \frac{\partial^2 T^*}{\partial y^2} = 0 \quad \text{_____(1)}$$

Is the elliptic partial differential equation (steady state) with two spatial dimensions, such that

$0 \le x \le a$ *and* $0 \le y \le b$ *where* $a = b = 1$

$T^*(x^*, y^*)$ is the steady state potential distribution in the given domain. The Dirichlet boundary conditions on three sides are homogeneous and on one side non-homogeneous are shown in Figure 1:

$$T^*(x,0) = 0, \quad T^*(x,b) = 75$$
$$T^*(0,y) = 0, \quad T^*(a,y) = 0$$
$$P(0) = 0 \ \ and \ \ P(1) = 1$$

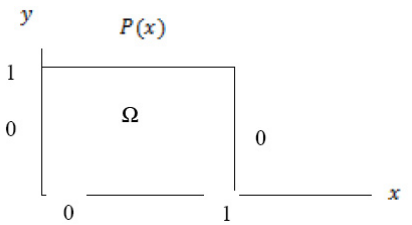


Figure 1. Rectangular domain of the assumed problem

The rectangular domain Ω is divided into finite number of square components. The division is such that each of the line and node of the field is shared with the connected elements other than the sides of the boundaries. The nodes and lines numbering are shown in Figure 2, as follows:

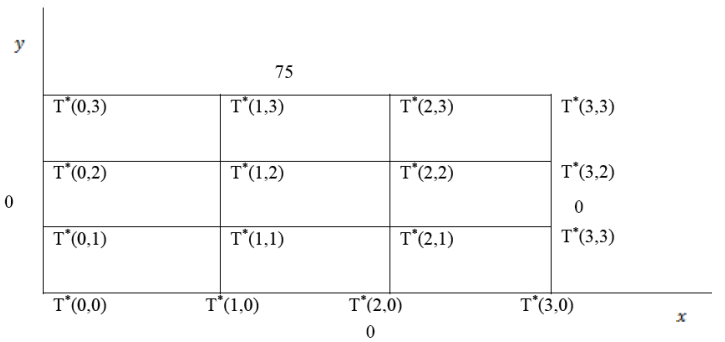


Figure 2. Potential distribution discretized region

III. The Finite Difference Methods (FDM)

FDM is a simple and easiest technique to numerical solutions of elliptic partial differential equations. In this problem approximated all the derivatives using finite differences. The discretization of the region in the directions $x^* and\ y^*$ of with a change of $\Delta^* x\ and\ \Delta^* y$ such that $\Delta^* x = \Delta^* y = h$ The discretized scheme is shown in Figure 3 below.

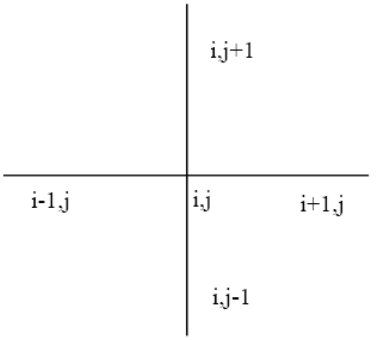


Figure 3. The discretized scheme in 2D

The equation to find the temperature at the particular nodes is

$$T^*_{i+1,j} + T^*_{i-1,j} + T^*_{i,j+1} + T^*_{i,j-1} - 4T^*_{i,j} = 0 \quad \text{_____(2)}$$

i. Derivation

Let us consider a Laplace Equation in two dimensional space on a rectangular shape like

$$T^*_{xx} + T^*_{yy} = 0 \quad \text{_____(3)}$$

With the conditions $0 \le x \le a \;\; and \;\; 0 \le y \le b \;\; a = b = 1$

The Dirichlet boundary conditions are

$$T^*(x,0) = 0, \quad T^*(x,1) = 75; 0 \le x \le 1$$
$$T^*(0,y) = 0, \quad T^*(1,y) = 0; \;\; 0 \le y \le 1$$
$$P(0) = 0 \;\; and \;\; pP(1) = 1$$

The grids are uniform in both $x \, and \, y$ directions. The objective is to find the approximate solution at the grid points only, that is $T^*_{ij} \approx T^*(x_i, y_j)$. The finite difference approximation to the partial derivatives at the grid point (x_i, y_j) are:

$$T^*_{xx} \approx \frac{T^*_{i-1,j} - 2T^*_{i,j} + T^*_{i+1,j}}{(\Delta x)^2} \quad \text{_____(4)}$$

$$T^*_{yy} \approx \frac{T^*_{i,j-1} - 2T^*_{i,j} + T^*_{i,j+1}}{(\Delta^* y)^2} \quad \text{_____(5)}$$

Plugging these approximations into the Laplace equation at the point, (x^*_i, y^*_j) we get

$$\frac{T^*_{i-1,j} - 2T^*_{i,j} + T^*_{i+1,j}}{(\Delta^* x)^2} + \frac{T^*_{i,j-1} - 2T^*_{i,j} + T^*_{i,j+1}}{(\Delta^* y)^2} = 0 \quad \text{_____(6)}$$

$$where\,(\Delta^* x)^2 = (\Delta^* y)^2 = h^2$$

$$T^*_{i+1,j} + T^*_{i-1,j} + T^*_{i,j+1} + T^*_{i,j-1} - 4T^*_{i,j} = 0 \quad \text{_____(7)}$$

Equation (7) is a discrete equation holds at every grid point (x_i, y_j) not on the boundary that is i = 1,2,3, - - - - - , *N and j = 1,2,3 - - - - -, N means (N-1) (N-1) equations*

The boundary conditions are:

$$T^*_{0,j} = 0, T^*_{N,j} = 0; 0 \le j \le N$$

$$T^*_{i,0} = 0, T^*_{i,N} = P(x_i),\ 0 \le x \le N$$

Note that the only interior points will be the unknowns. The equation for the temperature distribution at a particular node is:

$$T^*_{i,j} = \frac{T^*_{i+1,j} + T^*_{i-1,j} + T^*_{i,j+1} + T^*_{i,j-1}}{4} \quad \text{_____(8)}$$

The temperature at the four sides is given, at all the internal points the temperature is assumed. Here this study used the Gauss-Seidel iterative method for solving the system of equations. All the points, which have equal steps horizontally and vertically, the potential is distributed by the finite difference equation (8).

The gauss-Seidel iterative process for the numerical solution of the assumed problem is, shown in Table I:

TABLE I
NUMERICAL RESULTS

Iterations	$T^*_{1,1}$	$T^*_{2,1}$	$T^*_{1,2}$	$T^*_{2,2}$
01	0	0	18.75	18.75
02	4.6875	4.6875	23.4375	23.4375
03	7.03125	7.03125	25.78125	25.78125
04	8.203125	8.203125	26.953125	26.953125
05	8.7890625	8.7890625	27.5390625	27.5390625
06	9.08203125	9.08203125	27.83203125	27.83203125
07	9.228515625	9.228515625	27.978515625	27.978515625
08	9.3017578125	9.3017578125	28.0517578125	28.0517578125
09	9.3383789062	9.3383789062	28.0883789062	28.0883789062
10	9.3566894531	9.3566894531	28.1066894516	28.1066894516

The stopping criterion for the iterations is:
So at the 9th iteration

$$\left|\varepsilon_{i,j}\right| = \left|\frac{T^*_{i,j}\,present - T^*_{i,j}\,previous}{T^*_{i,j}\,present}\right|$$

$$\left|\varepsilon_{i,j}\right| = \left|\frac{9.3383789062 - 9.3017578125}{9.3383789062}\right| * 100 = 0.3921\% \approx 0.4\%$$

At the 10th iteration

$$\left|\varepsilon_{i,j}\right| = \left|\frac{9.4566894531 - 9.3383789062}{9.3566894531}\right| * 100 = 0.0196\% \approx 0.02\%$$

this is negligible so we stop and this solution is appropriate and reliable approximate solution.

IV. Analytical Solution

The two dimensions Laplacian equation with the boundary conditions are:

$$\nabla^2 T = 0 \_\_\_\_\_ (9)$$

BC:
$$T^*(0,y) = 0 \ , \ T^*(a,y) = 0$$
$$T^*(x,0) = 0 \ , \ T^*(x,b) = 75$$

Separation of variables is used to reduce PDE to ODEs. In the method of variable separation, it tried to find the solution in the form of product later on replaced the solution into Laplacian equation. The constant of separation is introduced by λ. Here only certain values of λ are allowable. The determined solution of Laplacian equation is only satisfied the boundary conditions at $\lambda < 0$. $T^*(x,y) = X(x)Y(y)$ these solutions have

$$X(x) = C_1 \cos \alpha x^* + C_2 \sin \alpha x^* \_\_\_\_\_ (10)$$

$u \sin g \ \ boudary \ \ conditions$

$$X(x) = C_n \sin(\frac{n\pi}{a})x^* \ ; \alpha = \frac{n\pi}{a} \_\_\_\_\_ (11)$$

$$Y(y) = A_n e^{\frac{n\pi}{a}y^*} + B_n e^{-\frac{n\pi}{a}y^*} \_\_\_\_\_ (12)$$

The product solution of Laplace equation with the given boundary conditions is shown below and represented in Table II:

$$75 = \sum_{n=1}^{\infty} 2A_n \sin(\frac{n\pi}{a})x \sinh n\pi \quad \text{_____(13)}$$

$$\int_0^a \sin(\frac{n\pi}{a})x \sin(\frac{m\pi}{a})x\,dx = \{0, n \neq m \text{ and } \frac{a}{2}, n = m$$

finally we get

$$T(x,y) = \frac{300}{\pi} \sum_{k=1}^{\infty} \frac{\sin(n\pi/a)x \sinh(n\pi/a)y}{n \sinh(n\pi)} \quad \text{_____(14)}$$

when n is odd that is $n = 2k-1, \quad k \in z$

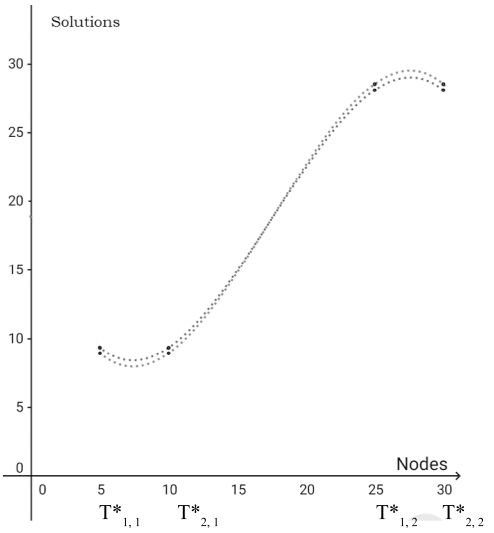
TABLE II
EXACT SOLUTION

Nodes	Exact Solution
$T_{1,1}$	8.94613
$T_{2,1}$	8.94613
$T_{1,2}$	28.5539
$T_{2,2}$	28.5539

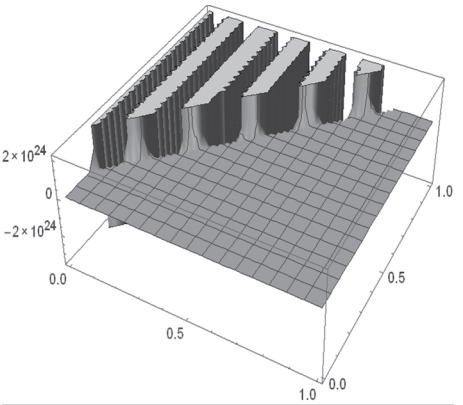
This is the exact solution of 2D-Laplace equation obtained by using variable separation with Dirichlet boundary conditions.

A. Graphical Representation of the FDM Solution and Exact Solution

Following graph a and b are the representation of the FDM solution:



Graph a: Comparison of Numerical FDM and Exact Result



Group b: Potential Distribution 3D Representation

V. Conclusion and Discussion

This study focused on the software uses in varies domains to obtain solutions numerically by using numerical methods particularly Finite Difference method (FDM) to solve 2D Laplace equations with Dirichlet boundary conditions. The graph and the tables I and table II results shown the good agreement of the exact and numerical solution obtained by FDM. Finite difference method is actual an average discretized domain method so it is more appropriate method as compare to other numerical methods for potential distribution. The results of table I showed that for the prediction of potential distribution in the regular rectangle domains Finite difference technique is superior in both competence and accuracy. To understand other quantities like potential distributions on irregular domains or flow and velocity distribution in varies geometries are the research areas in future for the researchers.

References

[1] M.L.Dhumal and S.B.Kiwne(2014). "Finite Difference Method for Laplace Equation",International Journal of Statistics and Mathematics,Volume 9(pp 11-13).

[2] Parag V.Patil and Dr.J.S.V.R.Prasad(2013). "Numerical Solutions for Two Dimensional Laplace Equation with Dirichlet Boundary Conditions",IOSR Journal of Mathematics, Volume 6(pp 66-75).

[3] Qian Z., Fu C.L., Li Z.P(2007). "Two regularization methods for a Cauchy problem for the Laplace equation,"Mathematical Analysis and Applications,May 2007.

[4] Morales M., Rodolfo Diaz R.A.,Herrera W. J.(2015) "Solutions of Laplace's equation with simple boundary conditions and their applications for capacitors with multiple symmetries",Journal of electrostatics, 78(pp 31-45).

[5] Lesnic D., Elliott L. and Ingham D.B.(1997) "An iterative boundary element method for solving numerically the Cauchy problem for the Laplace equation", Engineering Analysis with Boundary Elements, Volume 20.

[6] Li Z.C., Zhang L.P., Wei Y., Lee M.G., Chiang J.Y.(2015) "Boundary methods for Dirichlet problems of Laplace equation in elliptic domains with elliptic holes", Engineering Analysis with Boundary Elements, Volume 61.

[7] Lee M.G., Li Z.C., Huang H.T., Chiang J.Y.(2015). "Neumann problems of Laplace's equation in circular domains with

circular holes by methods of field equations", Engineering Analysis with Boundary Elements, Volume 51.

[8] Smith G.D. (1978). "Numerical solution of Partial Differential Equations 2nd Edition", Oxford, London.

[9] Ames W.F. (1992) "Numerical Methods for Partial Differential Equations 3rd Edition", Academic Press, San Diego.

[10] Lapidus L and Pinder G.F. (1982)"Numerical solution of partial differential equations in Science and Engineering", Wiley, New York.

[11] Greenspan D. and Parter S.V.(1965) "Mildly Nonlinear Elliptic Partial Differential Equations and their Numerical solution-II", Numerische Mathematik, 7(129-146).

[12] J.Thomas Beale and Anita T. Layton (2006). "On The Accuracy of Finite Difference Methods for Elliptic Problems With Interfaces", COMM.APP.MATH.AND COMP.SCL.VOL.1 NO 1.

[13] P.Jensen(1972). "Finite Difference Techniques for Variable Grids", Computers & Structures, 2(17-29).

[14] N. Perrone, R Kaos(1975). "A General Finite Difference Method for Arbitrary Meshes", Computers & Structures, 5 (pp 45-58).

[15] P.G.Martinsson(2013). "A direct solver for variable coefficient elliptic PDE's discretized via a composite spectral collocation method", Journal of Computational Physics, 242(pp 460-479).

[16] W.F.Ames(1997). "Numerical Methods for Partial Differential Equations, Academic Press, New York.

[17] A. Bueno-Orovio, V. M.Perez-Garcia, F.H. Fenton(2006). "Spectral methods for Partial Differential Equations in Irregular Domaim. The Spectral smoothed Boundary Method", SIAM, J.Sci.comput,28(3) (pp-886-900).

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Treatment of Reactive Routing Protocols Using Second Chance Based on Malicious behavior of Nodes in MANETS

[1*]M. R. Belgaum., [1]S. Soomro, [1]Z. Alansari, Member IEEE and [2,3]M. Alam
[1] College of Computer Science, AMA International University, Kingdom of Bahrain.
[2]Institute of Business and Management (IoBM), Korangi Creek, Karachi, Pakistan
[3]Universiti Kuala Lumpur, (IPSUniKL), Kuala Lumpur, Malaysia

***Abstract*-Mobile nodes of various routing protocols in Mobile Ad hoc Networks follow different strategies in transmission and receiving of data. Security, packet delivery and routing overhead are important concerns for any protocol during designing them. The presence and absence of malicious nodes in the network affect a lot on the performance of the protocol. This research focused on the study of the threats, attacks and reasons for malicious behavior of nodes in the network for reactive routing protocols in MANETS. DSR and AODV are the two reactive routing protocols that considered the study to propose a second chance strategy to given to the nodes considering the reason for malicious behavior to improve the packet delivery ratio and reduce the routing overhead in the network. A simulative study has conducted using Ad hoc Simulator (ASIM) considering the DSR and AODV routing protocols in the presence of malicious nodes and in the absence of malicious nodes that showed, that the packet delivery ratio is low and routing overhead is high in the absence of malicious nodes. The second chance strategy proposed considers the reasons for malicious behavior and helps the node to be reintegrate in the network to improve the packet delivery ratio and reduce the routing overhead.**

Keywords: MANETS, DSR Protocol, AODV Protocol, Malicious.

I. Introduction

A set of mobile nodes that perform basic networking functions in an infrastructure less environment said be a mobile ad hoc network (MANET). In the network, each node helps another node in forwarding or routing packets to the destination. Maintaining security is an important function of any of the routing protocol in each phase of the networking function [1]. Because of the non-static topological behavior of the network and due to being, the network open that allows the nodes to be add and deleted from the network anytime gives chance for the intruder nodes to disturb the normal routing process. In addition, if there does not exist a common regulatory authority for authenticating and guaranteeing the nodes then a reliable transmission is not possible.

* bmdriyaz@amaiu.edu.bh

A. Security Related Goals and Challenges

Security services are needed to make sure that the data is transferred over the network with reliability and the keeping the resources of the system protected. To attain the objectives, the categorizations of security services are availability, confidentiality, authentication, integrity and non-repudiation [2, 3].

1) *Availability:* Though the system is suffering from various problems like with bandwidth, connectivity but the availability service ensures that still the resources are available in a timely manner. Availability of a network can prove to harmful for example when the packets dropped and by resource depletion attacks.
2) *Confidentiality:* The information prevailing in the network is not to share among all unauthorized nodes and this achieved by Confidentiality. In order to achieve Confidentiality many encryption techniques can used to make only the authorized nodes can share the transmission of information and the private and public keys.
3) *Authenticity:* To prove a node as a legitimate user the network service used is Authenticity. The absence of this service can make any node in the network impersonate any node, and then having a total control capture and control over the complete network.
4) *Integrity:* The data which is been transmitted in the network can be modified either wontedly or sometimes unwontedly. The Integrity network service ensures that the information which is been transmitted is not modified.
5) *Non-Repudiation:* This service guarantees that the message transmission has done between the two parties and it cannot be denied. Also using this service, it helps in detecting and isolating of compromised nodes in the network.

Communicating through the network in safe and secure way has been a challenging task because:

1) Not being a stable infrastructure.
2) The links in the network are prone to break and not secure.
3) Scarcity or overload on the system resources.
4) The network topology being dynamic.

A study of the various on demand routing protocols in MANETS has done to analyze the threats and types of attacks in the routing protocols along with the reasons for the nodes to be malicious. The proposed approach will use to reintegrate the node back into network, which were previously as malicious.

II. Review Of Literature

To have a secure transmission various secure routing protocols have been designed for MANETS [4, 5, 6, 7, 8]. The features and functionality of the routing protocols studied from [9]. Based on their type and the functions they perform; the protocols categorized as follows.

TABLE I
Ad hoc Routing Protocols

Category	Type	Name of protocol
Flat Routing	Proactive	OLSR
		DSDV
		WRP
	Proactive	DSR
		AODV
		TORA
		ABR
		SSR
Hierarchical Routing		HSR
		CGSR
		ZRP
		LANMAR
Geographic Position assisted Routing		GEOCAST
		LAR
		DREAM
		GPSR

A. *Reactive Routing Protocols*

1) *Dynamic Source Routing Protocol:* The Dynamic Source Routing protocol a reactive routing protocol has two phases namely route discovery and route maintenance [10]. Initially the routes discovered for transmitting the packets between source node and the target node. A route cache maintained having the information of the recently used routes. In the route maintenance phase, it ensures that the packet safely transmitted to the target, as there is a chance of routes being broken due to the dynamic nature of the topology. More over the researcher simulated the protocol using NS-2 to focus on the performance of this protocol using the metrics like packet delivery ratio in the presence of malicious nodes and in the absence of the malicious nodes. It found that the packet delivery ratio in the presence of malicious nodes is more than in absence of malicious nodes.

2) *Ad hoc On Demand Distance Vector Protocol:* The AODV protocol explains its functionality [11]. It stated that the features of both DSR and DSDV combined. The author explains the working of AODV protocol along with the Black hole attack and its effect on the protocol. Finally, a Counter Algorithm proposed to prevent the Black hole attack on make the AODV Protocol secure. Metrics like Packet Delivery Ratio and Average End-to-End Delay used to measure the performance of the protocol. A method to identify the malicious node explained in order to avoid forwarding of the information to the malicious node in the

routing table. The solution given did not impose any overhead on the nodes in the network.

3) *Temporally Ordered Routing Algorithm:* The Temporally ordered Routing Algorithm works using the link reversal concept and this algorithm ensures that it results in no loops [12]. The functionality of this protocol works in three phases as follows: (a) Phase – I create the route initially (b) Phase – II concentrates on the maintenance of the routes and (c) Phase – III deletes the routes that are proved to be invalid. TORA protocol compared with other routing protocols like DSDV, OLSR with the metrics like Control Overhead, Packet Delivery Ratio, End-to-End Delay and throughput that it is better in performance with varying number of nodes using NS-2 simulator.

4) *Associativity Based Routing:* The Associativity Based Routing protocol does not have loops, free of deadlock and no duplicate packets [13]. It focuses on route longevity. As there are very few broken communication links and less need for reconstruction of the routes, the overhead involved is less. An improved version of ABR was to optimize the bandwidth and demand to reduce the overhead based on the position information was proposed. It concluded that the path setup time was long for the routes, which gave a scope for the future research to improve the ABR Protocol.

5) *Signal Stability-based Adaptive Routing Protocol:* The working of SSR routing protocol states that the large routing tables are not required for routing [14]. The network will not be congested with the control messages but a type of denial of service attack is a threat to this protocol. The Signal Stability Table maintained that has information of signal's strength of all nearby nodes. The protocol simulated in OmNet and a metric known as CPU usage considered measuring the performance. It proved that in the presence of malicious node the usage of CPU was more than in the absence of malicious nodes.

B. Types of Attacks

To have a secure transmission various secure routing protocols have been designed for MANETS [4, 5, 6, 7, 8]. The features and functionality of the routing protocols studied from [9]. Based on their type and the functions they perform; the protocols categorized as follows.

Attacks broadly classified into 2 categories as Passive attacks and Active attacks based on whether the operation in the network is disturbed or not [15].

Passive attacks allow the unauthorized node to snoop the data without altering it. The main target of the intruder node is to steal any information that being transmitted which in turn makes the confidentiality to fail. These types of attacks are very difficult to detect but can be prevented by using powerful encryption mechanisms.

Active attacks modify or even delete the data, which transmitted through the network. The effect of these types of attacks is that it brings down the performance of the network. The following is a brief description of various active attacks.

1) *Sybil Attack:* In this type of attack the intruder node pretends to have multiple identities and modifies and alters the information by directly communicating with the legitimate nodes [16].
2) *Resource Consumption Attack:* A malicious node in this attack always consumes the system resources like the bandwidth, battery power, memory, access points etc. A form of replication of packets makes the nodes to consume bandwidth also the battery power.
3) *Black Hole attack:* The intruder node in this attack announces that to reach any other node, it has the nearest path in the network. Whenever the intruder node receives a request to reach the destination, these node replies saying that it has the shortest route. If this reply message reaches to the node, which initiated the request earlier than the actual node, then the route is established and hence the intruder node has found the place between two actual communicating nodes causing denial of service behaving as man in the middle.
4) *Worm Hole attack:* The attacker in the worm hole attack broadcasts to the other nodes as a neighbor after hearing transmission of packet from some node. The malicious node now forwards the packet to another location creating a wormhole. These types of attacks prove to be very dangerous [17].

III. Methodology

The main aim of this research is to identify the various characteristics and approaches of the reactive routing protocols and study the attacks and the reasons for malicious behavior of nodes in reactive routing protocols in order to propose effectively a general framework for the same. The study aims to consider the reactive routing protocols with different types of attacks on them.

The study addresses the following questions:

1) Reasons for the malicious behavior of the nodes.
2) The effect of having malicious nodes in the network and eliminating malicious nodes in the network.

After the reasons for malicious behavior of the nodes are studied, the process of treating such nodes following a common strategy can be developed, which may help in maximizing the throughput and minimizing the routing overhead answered.

The study mainly uses the following reactive routing protocols:

1) DSR (Dynamic Source Routing Protocol).
2) AODV (Ad hoc On Demand Distance Vector Routing Protocol).

Factors Considered are:

1) Throughput.

2) Average delay

3) Routing overhead

A simulative and analytical research methodology has adopted in conducting the study. The various routing protocols have their own methodologies to send the information to the destination. This research considers the reasons for eliminating the nodes from the path due to the malicious behavior. The plan is to propose a strategy, which will help in maximizing the throughput and minimizing the routing overhead on the routing protocols and thus help in the selection of the most optimal routing path for any protocol to send the information to the destination.

The research organized as:

1) Study the reactive routing protocols.

2) Analyze the reasons for malicious behavior of the nodes.

3) Propose a strategy to handle malicious behavior.

IV. Discussion

The prominent objective of any ad hoc routing protocol is to have a safe and secure transmission which has become a challenging task because of (1) lack of security during communication, (2) changes in the network structure or topology, (3) resource constraints (e.g. battery power, bandwidth, memory, and CPU processing capacity) and (4) movement of node that makes the network topology dynamic. Most of the traditional routing protocols design could not take into account the aspect of security.

The reasons for malicious behavior of the nodes are studied and protection can be provided to keep secured and to handle these types of misbehavior like not forwarding the data or messages, though there is a secure and safe route traffic is deviated, though there are no error packets are rerouted, not observing the errors and lack of sending error messages, or vice versa, frequently updating the routes unusually.

One of the malicious behaviors like not forwarding the data or messages considered in this research. Again, there could be many reasons for a node not to forward the messages or data like Nodes failure, link and network failures, transmission energy and power failures.

Simulations implemented using Ad hoc Simulator (ASIM), a simulator that developed to simulate protocols in MANETS. ASIM developed for wireless ad hoc networks. The Physical Terrain Dimensions considered is 500m X 550m. The range of communication is from 100m to 200m. Considering the number of nodes from 30 to 50 to be varying due to dynamic topology in a

network, the speed is uniformly distributed between 0 and 20 m/s to offer a range of users that are in a fixed location or in motion. In this simulation, nodes move according to the random waypoint mobility model. The data communication pattern in this study uses multiple source destination pairs, each sending a constant Bit Rate (CBR) flow of 4 data packets per second. The simulation time chosen to be long enough to potentially roam the whole area. The Application data packet size is 512 bytes. The factors varied are the total number of nodes in the network, the transmission time, Communication Range and the different network topologies.

Metrics like Packet Delivery Ratio and Routing Overhead considered both in presence of malicious nodes and in absence of malicious nodes to compare the performance of reactive, the routing protocols selected in this research.

Packet Delivery Ratio = Total no. of packets delivered / Total no. of packets sent.

Overhead = Total no. of routing messages transmitted / Total no. of routing messages received.

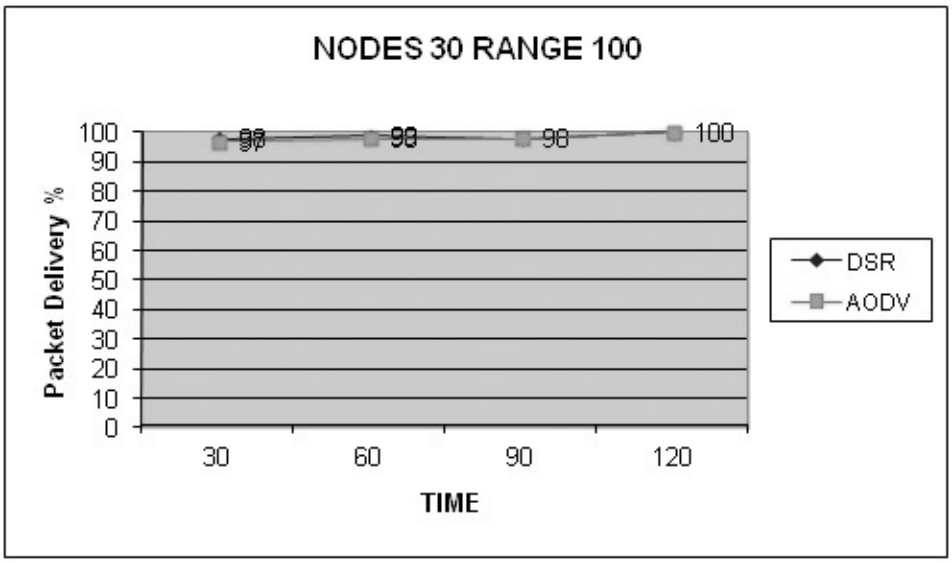


Figure 1. Packet Delivery Ratio in the presence of malicious nodes.

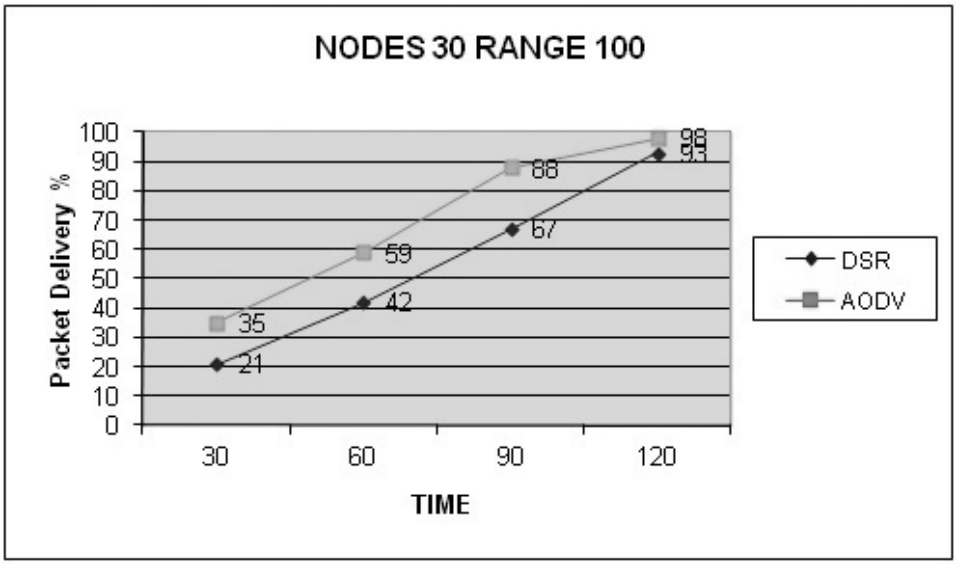


Figure 2. Packet Delivery Ratio in absence of malicious nodes.

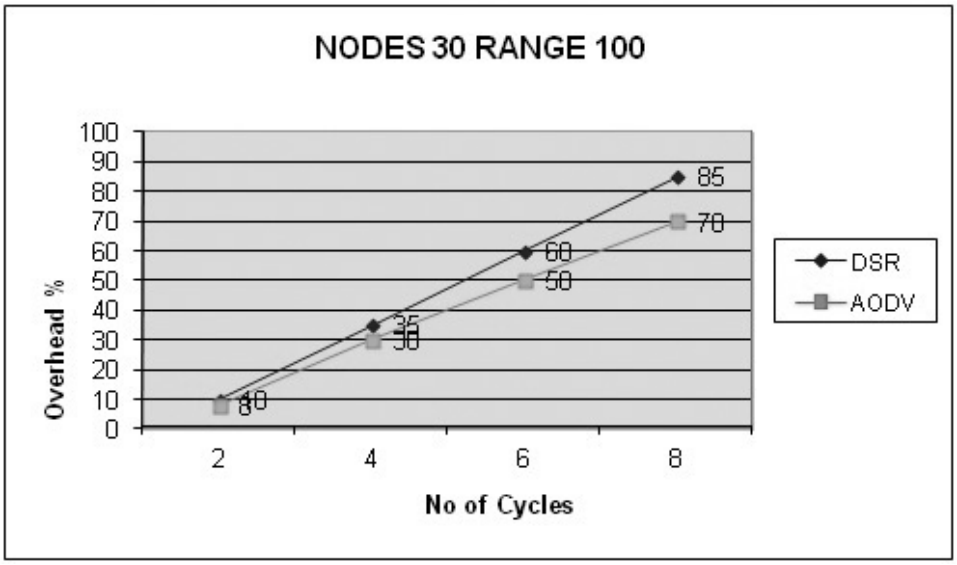


Figure 3. Overhead ratio in the presence of malicious nodes.

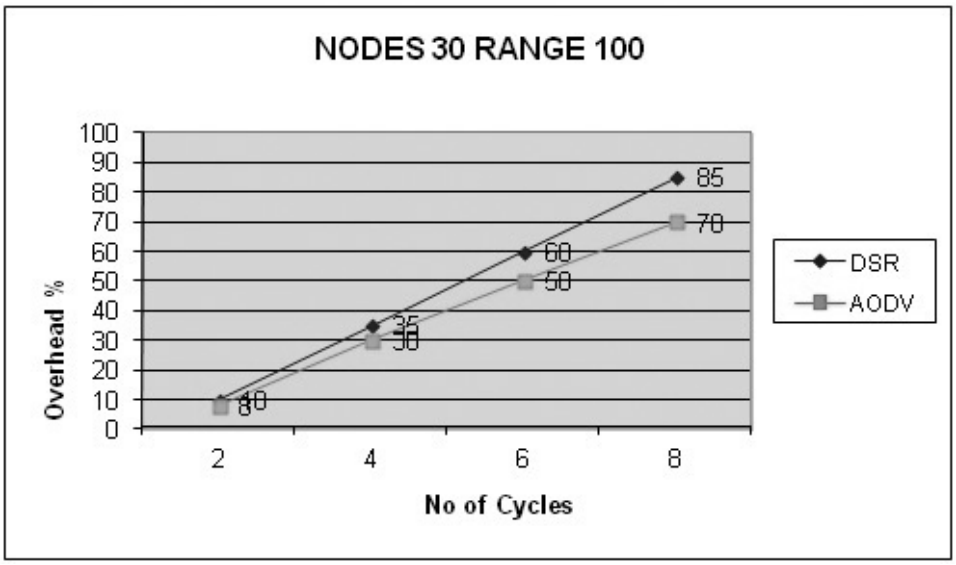


Figure 4. Overhead ratio in the absence of malicious nodes.

From the above figure 1 and figure 2, it is very clear that the packet delivery ratio is high in the presence of malicious nodes and is low in the absence of malicious nodes. When compared with the work done by other researchers where in their approach resulted in reduced packet loss in presence of malicious nodes in Trusted AODV (TAODV) protocol [18]. The results obtained here show more throughput in both DSR and AODV protocols with our proposed approach. Moreover, from figure 3 and figure 4, it is evident that the routing overhead is increases when there are no malicious nodes as all the malicious nodes eliminated from the path if they found to be misbehaving in DSR and AODV protocols. Therefore, elimination of nodes results in less number of paths to reach the destination and therefore increases the overhead. In one of the similar kind of research, only DSR protocol considered in eliminating the malicious nodes, which increases the overhead on the protocol in finding safe route to reach the destination [19]. Now in order to have high packet delivery ratio and to reduce the overhead, an approach of second chance proposed here for each of the node based on trust and previous behavior. In other words, instead of simply eliminating the node from the path, the reasons for its misbehavior could studied and each node could give a chance to be reintegrate into the network based on the reason how it misbehaved. The nodes after getting a second chance get added in the network will prove to be fruitful as there will be more number of paths available resulting in high packet delivery ratio and reduced overhead as there will be less path rejections.

```
if (node is malicious)
{
check the trust value and check past behavior;
if (within tolerance)
    {
     forward data packet to next hop;
     update trust;
         }
else
  {
eliminate the node from the path;
       }
 else
{
forward data packet to best neighbor1;
}
}
```

V. Conclusion

The proposed approach intends to improve the packet delivery ratio and reduce the routing overhead caused in the absence of malicious node due to frequent topological changes in ad hoc networks. A second chance strategy for the reactive routing protocols like DSR and AODV proposed to make a node reintegrated into the network considering the reasons for misbehavior of the nodes. A simulation analysis done between DSR and AODV protocols in both the presence and absence of malicious nodes using the metrics packet delivery ratio and routing overhead to propose the second chance strategy for the node based on the trust rating. Using this strategy, it will improve the packet delivery ratio as there will be more paths available to reach the destination and routing overhead will reduced as there will be less path rejections. In future changes in these protocols can done by using this approach and a simulative research can done.

References

[1] Ewa Niewiadomska-Szynkiewicz · Andrzej Sikora · Joanna Kołodziej, "Modeling Mobility in Cooperative Ad Hoc Networks," Springer-Mobile Networks and Applications, Vol 18, No. 5, pp. 610-621, 2013.

[2] A. L. Sandoval Orozco, J. Garcıa Matesanz, L. J. Garcıa Villalba, J. D. Marquez Dıaz, and T.-H. Kim, "Security Issues in Mobile Ad Hoc Networks", International Journal of Distributed Sensor Networks, Volume.8,No. 12, 2012

[3] Christos Xenakis, Christoforos Panos, Ioannis Stavrakakis, "A comparative evaluation of intusion detection architectures for mobile ad hoc networks", Vol. 30, Issue 1, pp 63-80, 2011.

[4] C. S. R. Murthy and B. S. Manoj, "Ad Hoc Wireless Networks: Architectures and Protocols", Prentice Hall PTR, 2004.

[5] Azzedine Boukerch,Begumhan Turgut,Nevin Aydin,Mohammad Z. Ahmad,Ladislau Bölöni ,Damla Turgut, "Routing protocols in ad hoc networks: A survey" -Science Direct Volume 55 issue 13. 15, pp 3032–3080, 2011.

[6] Said EL KAFHALI, Abdelkrim HAQIQ, "Effect of Mobility and Traffic Models on the energy consumption in MANET Routing Protocols." International Journal of Soft Computing and Engineering, Vol. 3, No. 1, pp 242-249, 2013.

[7] Mahfuzur Rahman Bosunia, daniel P. Jeong, Chanhong Park, Seong-Ho Jeong, "A New Routing Protocol with high energy effcient and reliability for data delivery in Mobile Ad Hoc Networks", International Journal of Distributed Sensor Network, Vol.11, No. 8, 2015.

[8] Gaurikar Talwar, Hemika Narag, Kavita Pandey and Pakhi Singhal, "Analysis of different Mobility Models for Ad hoc On-Demand Distance Vector Routing Protocol and Dynamic Sourse Routing Protocol", In Proceedings of Fourth International Conference on Networks and Communications(NetCom) by Springer, 2012.

[9] Muthana Najim Abdulleh, Salman Yussof, Hothefa Shaker Jassim, "Comparative study of Proactive, Reactive and Geograhpical MANET Routing Protocols", Journal of Communications and Network, Vol.7, No.2, pp.125-137, 2015.

[10] Zeyad Ghaleb Al-Mekhlafi, Rosilah Hassan, "Evaluation study on routing information protocol and dynamic source routing in Ad-Hoc network", IEEE 7th International Conference on Information Technology in Asia(CITA 11), 2011.

[11] Yaser khamayseh, Abdulraheem Bader, Wail Mardini, and Muneer BaniYasein. "A New protocol for Detecting Black Hole Nides in Ad Hoc Networks," International Journal of Communication Networks and Information Security, Vol. 3 Issue 1, pp.36-47, 2011.

[12] Tamilarasan-Santhamurthy, "A Quantitative study and Comparison of AODV, OLSR & TORA routing protocol in MANET," International Journal of Computer Science Issues, Vol.9, Issue 1,364-369, 2012.

[13] Fawaz.A.M.Masoud, Said Abu Shaar, Ayman Murad and Ghassan Kanaan, "Enhanced Route-Reconstruction Method for ABR protocol for MANET," Journal of Computer Science 2(12) : 859-871, 2006.

[14] Yufeng Chen, Zhengtao Xiang, Yabo Dong, "Multi-Fractal Characteristics of Mobile Node's Traffic in Wireless Mesh Network with AODV and DSDV Routing protocols," Wireless pers Communications, 741-757, 2011.

[15] T. Navaneethan, M. Lalli, "Security attacks in Mobile Ad hoc Networks – A Literature Survey," International Jouranl of Computer Science and Mobile Applications, Vol 2, Issue 4, pp 1-7, 2014.

[16] Akhila.V, Meghashri.M.E, "Preventing Sybil attack in Ad-Hoc network using Secret Key technique," International Journal of Computer Science and Information technology, Vol.3, Issue.3, pp-325-330, 2015.

[17] Haiyan Liu1, Zhanlei Shang, "Comparing the Performance of the Ad Hoc Network under Attacks on Different Routing protocol", International Journal of Security and its Applications, Vol.9, No.6, pp. 195-208, 2015.

[18] Anil Kumar Gupta, Deepti Mehrotra, "Detecting and Dealing with malicious Nodes Problem in MANET," International Journal of Scientific and Engineering Research,Volume.4,Issue 7, pp.161-166, 2013.

[19] Yaser Khamayseh, Ruba Al-Salah, Muneer bani Yassein, "Malicious Nodes Detection in MANETS:Behavioral Analysis Approach," Journal of Networks, Volume 7, No.1, pp.116-125, 2012.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Model and Technique over Software Requirement Prioritization

[1*]Khurram Ejaz, [2]Ayesha Amjad
[1]Universiti of Teknologi Malaysia (UTM), Johor, Malaysia
[2]National University of Computer and Emerging Sciences (FAST), Islamabad, Pakistan

***Abstract*- Requirement prioritization play a significant part in overcoming problems related to requirements and it use to increase customer satisfaction. Requirement prioritization is use to verify the correct functionality of product and guarantee that the software is built within the given constrains, like budget, cost, value, time and etc. Requirements prioritization reliant on the specific requirements of customer along with prediction of importance and cost of each requirements. The approaches proposed in modern days to prioritize requirements have not been widely used because of its complexity, its inconsistency and its time-consuming factor. This paper represents an analysis on obtainable prioritization techniques based on cost and benefit and drawbacks evaluation of requirements. This paper represents a new approach of requirement prioritization, which can be easily managed, implemented and used as decision-making device by decision maker for requirement prioritization. It also overcomes the drawback of existing requirement approaches. The results of newly proposed approach shows that proposed model for requirement prioritization is simple and is useful for more refined critical decisions of prioritization, keeping in view the cost and benefit.**

Keywords: Software engineering, Requirement prioritization, Requirement Engineering, Elicitation, Method approaches, Negotiation approaches, Ordinal scale, Systematic research, Greedy algorithm, cost, risk, requirement.

I. Introduction

Software engineering has become an important discipline for creating the software that full fill the customer needs. In requirement engineering in the elicitation phase the analyst collect information from the stakeholders about their requirements, remove ambiguity and identify problem faced by them in requirements. They try to find best solution and make plans to develop software. Decision makers in software organizations during development phase are facing different problems, in selecting correct set of requirements for a specific release. To select correct set of requirements for implementation they must know the priority of every requirement. One of the characteristic of the requirement is that it can be explicitly prioritize. When expectations of customer are high, time given for development is less and resources are limited then it is difficult to carry out all requirements.

The requirement prioritization is useful for developer to take decisions about the implementation of requirements on the basis of many characteristics of requirements; such as requirement's benefit to user and business value, implementation cost, time etc. Industry analyst has developed several

* khurram.ejaz49@gmail.com

techniques for requirement prioritization that involve value and cost of individual requirement. Cost characteristic of the requirement help management of decision maker to use budget easily and in an efficient way. Mostly in requirement prioritization lower priority requirements are ignored so that their cost is helpful in achieving higher priority requirement to maximize business value. Sometimes requirement with greater priority is dependent on a requirement with lower priority, but to archive business benefits, decision makers are forced to implement lower priority requirements.

To gain customers satisfaction and business value the decision makers try to maximize the product value within given budget and time constrain. Until now none of existing companies know how to assign priorities to requirements. The main concept of requirements prioritization is the requirements prioritization process, which is consists of actions, that are performed on each requirement during their prioritization. The two main factors associated with each requirement are their benefit and cost. A procedure for prioritizing customers' requirements should be simple and fast, easy to manage and should give trustworthy results. By performing a survey on different requirement prioritization techniques a question was assumed that in which way requirements prioritization based on benefit and cost estimation currently is supported and by which available methods? Study answered it by using a systematic review of writing, drawing on our earlier work. The purpose of this study is to gain understanding of requirement prioritization techniques implanted in companies.

The prioritization method consists of three stages:

I. In preparation stage all the requirement from the customer are gained and arranged. Requirements are arranged in such a way that decision maker can apply required prioritization technique.

II. In execution stage prioritization of requirement is done using a specific requirement prioritization technique.

III. In presentation stage the optimized result of prioritization is presented.

II. Prioritization Techniques

In this section outline of different prioritization techniques and their drawback are described. The descriptions are defined one by one as followed.

A. Requirement prioritization techniques:

There are two category of prioritization.

I. Methods approach: methods are based on assigning quantitative values to different aspects of requirements.

II. Negotiation approach: it includes giving priorities to requirement according to agreement between stakeholders.

The requirements uncertainty prioritization approach (RUPA) [1] is basic prioritization technique in which Numerical Assignment is shaped as Extensive Numerical Assignment by possibility distribution and grade intervals. The main idea of transformation of Numerical Assignment was to have set of inputs, which are flexible enough to lodge vagueness. The backbone of the approach is interval evidential reasoning algorithm, which is used to combine the vague assessments of stakeholders. Framework consist of identifying the group of stakeholders, which should include developers to provide cost and risk ratings, customers who provide benefit, penalty ratings and Project Manager who judge conflicts and makes trade-off decisions. RUPA was implemented by Indian University named as Adikavi Nannaya. It is simple and fast prioritization technique but it does not give correct and trustworthy results and it is not flexible enough to lodge any ambiguity.

Numerical assignment is an approach that is use to distribute the requirements into different categories based on their importance like high, medium and low. It is important that each group help the stakeholders to do consistent classification of requirements. The group can vary but mostly groups are critical, standard and optimal. It prioritizes the requirements on an ordinal scale. 11 companies implemented different prioritization techniques. Out of 11 companies' majority used numerical assignment to prioritize their customer requirements. The sets of categories used by companies was critical, medium, and not critical. The main drawback of numerical assignment is that the usefulness of the requirement priorities is finished because the stakeholders are forced to divide the requirement into three groups.

The Detection and prevention (DDP) [6] method help developers to select the development activities (such as procedures, tests, and analyses) cost effectively and it prevent the hardware defects and detect. According to framework of DDP, assign each requirement a weight, which help developers use these to determine the importance of requirement. Then risks relate to requirements are assigned to indicate each delay is caused in customer requirement. Each mitigation is assigned an effort a cost. The cost is generally the financial cost, but other factors such as resources, schedule and utilization of memory can be considered. DDP cannot deal with continuous requirements and cost is indirectly achieved.

Value-oriented prioritization (VOP) is prioritization an approach in which requirement influences the business values of an organization [7]. The framework of VOP helps the developers to recognize the core values of the business and the associations between those business values. An executive of a company classify the business values. It uses a simple scale to weight the values of requirements according to the importance of an organization. Its framework identifies the weight of business risk. VOP formulate a prioritization matrix using core values and business risks. VOP

was tested by TBI (Technology Builders Incorporated) they decided to bound the value to a suitable size. TBI got many benefits by using VOP then the company's leaders firmly believe that VOP contributed directly to TBI's success.

The 100-dollar test prioritization and ranking techniques [8] are almost same. In 100-dollar test prioritization the stakeholders are given 100 imaginary units (money, cost, benefit, hours, etc.) to distribute among the requirements. In ranking the requirements are placed in a rank. The most important requirement is placed on top of the rank at first position and the least important is placed at the bottom at rank n in the list. For n requirements the requirements will be prioritized into n ranks. The list of ranked requirements could be obtained in a variety of ways like using sorting algorithms example by using the bubble sort or binary search tree algorithms.

Cost–value approach [5] is an analytic technique. Which help the developers to level candidate requirements by using two dimensions: according to their value or importance to customer and users, and according to the estimated cost of implementation of the requirements. The cost–value method is an entrenched analytical technique, with reasonable effort, provides a clear indication of the costs and values of all candidate requirements. To examine candidate requirements, Analytic Hierarchy Process (AHP) was used in which comparison of requirements is done pair-wise according to their value and cost. Ericsson's Radio Access Network project (RAN) implemented cost value approach. They identified 14 requirements then applied AHP on them. The software managers were able to efficiently and accurately prioritize their requirements. In cost value approach there are several issues of interdependencies of requirements and by applying AHP there is lot of complexity.

Quality function development (QFD) is used as methodology to describe relationship between the customer requirements and technical attributes like cost, importance, value. A relationship matrix is made to show impact of technical attribute of requirement and correlate show the dependencies of the technical attributes of the requirements.

Goal skill preference is used to generate a software design. In an analysis phase of goal preference, the requirements are given as input and it generate a set of ranked alternative that fulfill the customer's need. In design phase of software, the set of alternatives is used as software component for its architecture. To remove the drawbacks of above mentioned prioritization technique a new technique is developed, which is simple and east to implement.

III. Parameterised Requirements Prioritisation

After conducting literature review this paper extracted three major parameters from nine papers. The names of parameters are requirement, cost and risk. These are the major finding and have scales all of parameter with:

- *1= good finding.*
- *2= Average finding.*
- *3=low findings.*

TABLE I
SUMMARY OF REQUIREMENT PRIORITIZATION MODELS

S No	Models	Cost	Risk	Requirement	Remarks on high scale
1	A model for making early requirement	1	3	2	Cost is efficient due to resources
2	Requirement handles IKIWIS cot and Rapid change	3	1	1	Risk=time Requirement=va able
3	Value oriented model	1	3	3	Requirement=va e
4	Prioritization based on cost and benefit co prediction	1 s	1	1	Cost=software economics Requirement=co
5	A cost value model.	1	3	3	Requirement=su et.
6	Prioritization o quality requirement	3	3	3	
7	Requirement prioritization	3	3	3	
8	Requirement	1	3	1	Requirement=pe
	uncertainty approach				alty Cost=optimal solution
9	Requirement in priority in agile	1	3	1	Cost=size Requirement =acquisition

After studied of different papers this study concluded that each author is trying to propose much complex/complicated or ambiguous model. Authors of this paper have suggested or proposed a very simple model for requirement prioritization and use more refined ones, especially when a more sensitive analysis is needed for critical decisions.

IV. MODEL OF PARAMETERISED REQUIREMENTS

In the proposed model the square shows entity and oval sign represent process. The whole scenario proposed software prioritization consist of two entities and five process and main process consist of requirement, cost and risk. Requirements are going to ask from client by the project manager; then the project manager will compile the requirements and get compile data; later that compile data is converted into optimal data. Optimal data will help to decide cost and on basis of cost will enable us to determine if project is on risk or in profit.

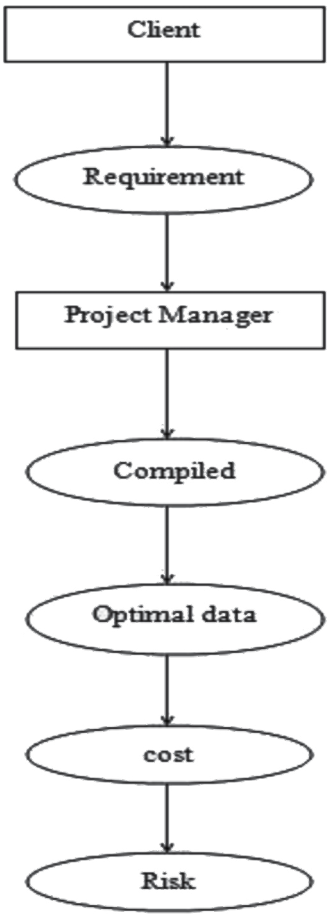


Figure 1: New proposed model for prioritization

A. Greedy algorithm (High level algorithm)

In general, greedy algorithms have five components:

1) A requirement set, from which a solution is created

2) A selection function, which chooses the best requirement to be added to the solution

3) A feasibility function, that is used to determine if an estimated requirement can be used to contribute to a solution

4) An objective function, which assigns a value to a solution, or a partial solution, and

5) A solution function, which will indicate when to discovered a complete solution

B. B. Low level algorithm

Low level algorithm, steps of the algorithm are given below:

1. *Find requirement from user side...*
2. *Requirement [ ]=a [n]*
3. *Estimated [] =Requirement []*
4. *While (True)*
5. *If (Estimated [n] == weight[n]);*
6. *Cost=Estimate*
7. *Print "Cost"*
8. *Else If (Estimated[n]>=weight[n])*
9. *Risk=Estimated*
10. *Print "Cost"*
11. *Else if (Estimated[n]<=weighted[n])*
12. *Cost=Estimated*
13. *Optimal solution*
14. *End*

In above algorithm first the project manager will ask the requirements from the customer are stored in an array. a [ ] is such alpha numeric array which will get the maximum requirements from user and this array in another array of Requirement so that algorithm will become user readable. Requirement [] containing possible array which tried to extract functional requirement from the possible requirement array so that it can lead this algorithm to successful system. Then in next step try to give ranks to requirements and then *1.* These ranks give weight to requirements. To give weight and have to initialize another array of weight. If the rank of requirement is high than it can find the cost of the requirement; that will be effective value, which leads towards a very successful system. If in this case weight of all requirements are highly ranked than best or optimal requirement cost will be able to find. But if weight requirements are less than estimated than system will be on Risk. The variable n which help to determine number of customer requirements.

V. Implementation of the Model

Flow chat describing the flow to requirement prioritization of proposed model, as shown in Figure 2.

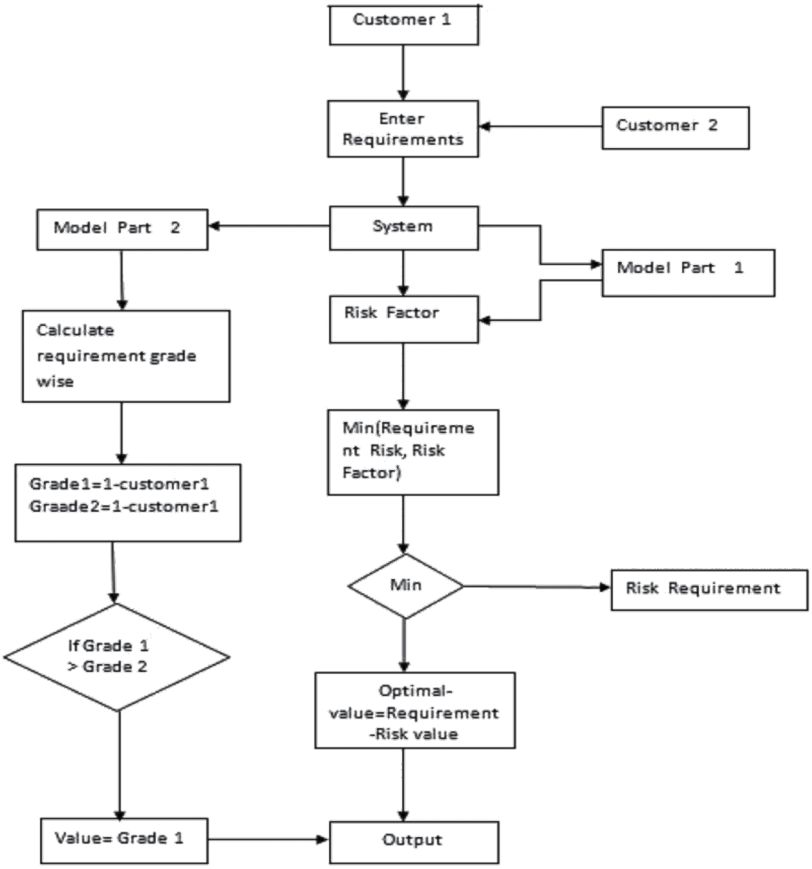


Figure 2: Flow chat of proposed model for prioritization

An experiment was performed by coding this model in Microsoft visual studio 2013 using C++ language. Few functions are given below:

Function for model part 1:

The function of model part 1 is given below:

```
int model part 1( )
{
float wt_act,wt_database, wt_stock, wt_model, wt_customer, total_wt, minimum, risk,req_
risk,value; if (choice2==1)
{
total_wt=wt_act+wt_database+wt_stock+wt_model+wt_
    customer; total_wt=total_wt/100; cout<<total_wt<<endl; float value1, value2,
```

```
                value3, min,risk_fact=0.1, mitg; value=100.0*(1.0-min(1.0,risk*req_risk));
                value1=risk*req_risk; value2=min(1.0, value1);
    if(value1<1.0)
        {
        min=value1; value2=min;
        }
    value3=1.0-value2; value=(100.0*value3)/100.0; minimum=min(wt_act,wt_database);
for(float i=r;i>=0;i--)
        {
        value=value-risk_fact; cout<<"mitigated value"<<value<<endl;
        }
        }
```

This novel model mainly consists of two entities named as customer and developers. These two entities have three main parameters, which are requirement, risk and cost. These parameters will be input in the model which will give some discrete values and these discrete values are point based like 0.1 to show probability. The main and important parameter is requirement. Customer will enter requirement and will see simulated values from various formulas. In this paper model is divided into two parts. In part I single Customer, the one who can enter the requirement. These requirements will pass into system. System will check the requirements, which will hold two things, risk factor and other is requirement factor. Our defined function will take minimum factor from these two parameters. In this system given weight age about 0.1 to general risk value from requirement and minimum requirement value will be subtracted from risk value for the optimal value.

Function for model part 2:

The function of model part 1 is given below :

```
float model part 2( )
{
    model1 l2,l3; float j,t,t1;
    int i;
    cout<<endl<<"enter first requirement"<<endl;
    t=l2.l1.req1; cin>>t;
    t=1-t;
    cout<<endl<<"enter second requirement"<<endl;
    cin>>l3.l2.req1; t1=l3.l2.req1; cin>>t1; t1=1-t1; cout<<t1; cout<<endl; float l4,k,g;
l4=t*t1;
```

```
cout<<endl<<"total of two stake holders probability"<<l4<<endl;
//k=1-l4; //cout<<k; if(t>=t1)
{
g=t;
cout<<"this is high priority requirement from two stake holder"<<g;
}
//l1.req[i]+l2.req[j];
//for(int i=0;i<=4;i++)
{ l2.req[i]=5*l1.req[i]; cout<<l2.req[i];
}
}
```

In second part of model, study proposed two stakeholders means two customers and they are giving requirements individually. In this model both customer 1 and customer 2 will enter the requirements in second part. After having requirements, will try to find grade value. Total probability is 1 value for customers 1 and customer 2. After subtracting requirement from total will find grade values. If grade value will be higher, then probability of getting requirement is important.

VI. Experimental Results

Here above mentioned model in figure 2 is followed, in which the diagram of proposed system is drawn, is dividing into two path, one is model part 1 and other is model part 2, in model part one mathematical formulation is drawn and can be shown in graphe1 below; in graph x axis is consisting of input and output parameters and y axis is giving values of input parameter, higher value from graph. It is noted that that higher values are mitigated higher risk values after taking five requirements value in this simulation, which can be extend up to n .

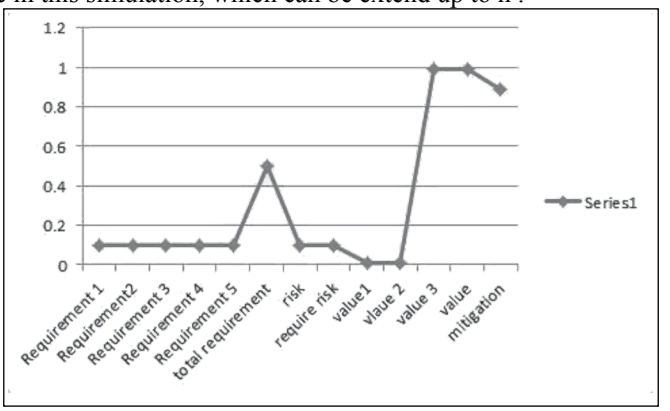


Figure 3: Graph for proposed model part 1

In model part 2 if consider the case, where two customers or they may be more than two from figure number 2; it can be clearly shown the simulation and also shown in the graph, which have been found from code below; where x axis is consisting of two customers and they are after giving their probability value and finding grades requirement value and customer grade parameter can be found on x axis and on y axis can see their values. Higher value shows that customer has high value of risk whereas low value shows less risk, from graph risk become 0.7 from 0.8 because using directly minimal value to minimize the risk.

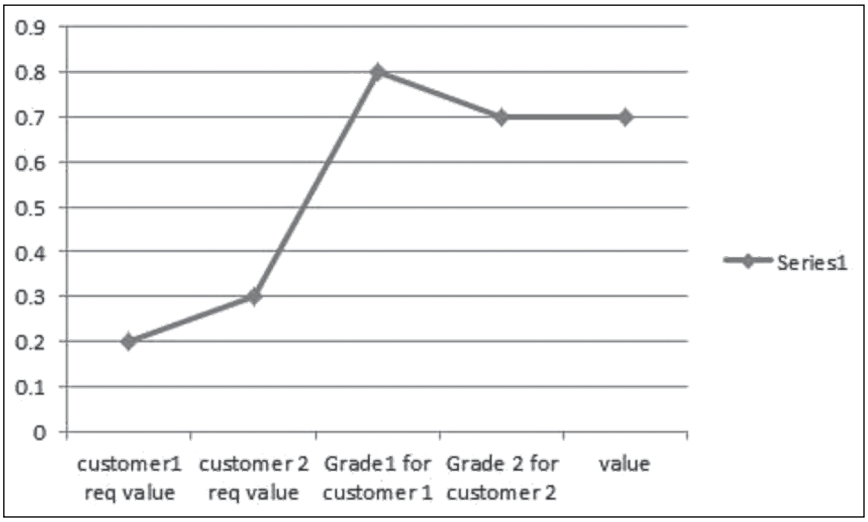


Figure 3: Graph for proposed model part 2

VII. Conclusion and Future Direction

Requirement prioritization is introduced to find most important requirement that add greatest value to business. It helps out the developer to take important decisions related to requirement implementation. The low priority requirements are dropped so that time, cost and resources are saved and Prioritization technique to be implemented must be fast, easy to manage and should give trust worthy results. The model proposed in this paper is simple and use more developed ones, when a more sensitive analysis is done for requirements. In proposed model the requirements are given by client to the project manager and the project manager will compile the requirements into compiled data and that compiled data will be converted to optimal data. Optimal data will help to determine cost and on basis of cost it will be enabled that either the project is on risk or in profit. The results show that proposed model for requirement prioritization is simple, fast and it use provide refined decisions for prioritization according to cost and benefits.

An increase in customer requirements causes an increase in requirements prioritization techniques. When there are many requirements then the prioritization could be done in early stages and need to manage with change in customer needs. Many new prioritization techniques are introduced in industry, but their scalability is not tested. Future efforts could be made to refine this model. More investigation could be done on the relationship between resources of requirement and the technical attributes. A survey on requirement prioritization techniques is performed.

This work is a model in such a way that system is involving three parameters requirement, risk, cost. Customer is main entity who is giving specification to developer and developer tries to accurate requirements so that a risk free system can delivered, but in future this may transplant into an intelligent system, which will consist of various models and this system will play the role of developer and customer will give requirement to intelligent system and intelligent system will apply intelligent calculation on user given requirements and find an optimal risk free customize solution on hands on.

References

[1] Persis Voola," Requirements Uncertainty Prioritization Approach: A Novel Approach for Requirements Prioritization ",September 2012.

[2] Abdel Ejnioui1, " A Simulation-Based Fuzzy Multiplet Attribute Decision Making for Prioritizing Software Requirements ",2012

[3] Svensson, R.B. ," Prioritization of quality requirements: State of practice in eleven companies ",September 2011.

[4] Racheva, Z. , " Do We Know Enough about Requirements Prioritization in Agile Projects: Insights from a Case Study" ,October 2010.

[5] Herrmann, " Requirements Prioritization Based on Benefit and Cost Prediction: An Agenda for Future Research ", September. 2008

[6] Feather M.S, "A Broad, Quantitative Model for Making Early Requirements Decisions ",22 February 2008 .

[7] Azar ,"Value-Oriented Requirements Prioritization in a Small Development Organization ", 08 January 2007 .

[8] Patrik Berander, " Requirements Prioritization",2005.

[9] Nadina Martínez Carod ,"Classifying Software Requirement Prioritization Approaches", 2005

[10] Lena Karlsson1,"Requirements Prioritization: An Experiment on Exhaustive Pair-Wise Comparisons versus Planning Game Partitioning", 25 May 2004 .

[11] Patrik Berander ,"Using Students as Subjects in Requirements Prioritization", 20 August 2004

[12] Boehm, B. "Requirements that handle IKIWISI, COTS, and rapid change ", August 2002 .

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Tributary Mapping Multiplexing an Efficient Technique for High Speed Fiber Optic Communication

[1*]U. Illahi, [1]J. Iqbal, [1]M. I. Sulaiman, [2,3]M. Alam and [4]M. S. Mazliham
[1]British Malaysian Institute (BMI), Universiti Kuala Lumpur, Gombak, Selangor, Malaysia
[2]Institute of Business and Management (IoBM), Korangi Creek, Karachi, Pakistan
[3]Universiti Kuala Lumpur, (IPSUniKL), Kuala Lumpur, Malaysia
[4]Malaysia France Institute (MFI), Universiti Kuala Lumpur, Bander Baru Bangi, Malaysia

***Abstract*- A novel technique of multiplexing called Tributary Mapping Multiplexing (TMM) is applied to a single channel wavelength division multiplexing system and performance is monitored on the basis of simulation results. To elaborate the performance of TMM in this paper, a 4-User TMM system over single wavelength channel is demonstrated. TMM showed significant tolerance against narrow optical filtering as compared to that of conventional TDM at the rate of 40 Gbit/s. The above calculations are made by optical filter bandwidth and dispersion tolerance that was allowed at minimum. The spectral efficiency achieved by this TMM was 1 b/s/Hz and it was executed by using transmitters and receivers of 10 Gbit/s without polarized multiplexing. The high spectral efficiency, high dispersion tolerance and tolerance against strong optical filtering makes TMM an efficient technique for High Speed Fiber Optic Communication.**

Keywords: Fiber Optic Communication, Tributary Mapping Multiplexing, Spectral Efficiency, Optical Filtering, Wave Length Division Multiplexing and Dispersion.

I. Introduction

The demand for high date rate is increasing day by day. The latest applications used these days require high capacity along with data rate, in this regards the researchers are working day and night to innovate new techniques, which can increase the capacity of optical fiber communication network. Bandwidth is major issue these days and efforts are being made to save the bandwidth by making its utilization better. The demand for high speed communication systems is increasing tremendously and an increment of 40 % per annum is observed. Fiber optic communication is expensive in the aspect that it's laying, operation and maintenance is costly. But on the other hand its capacity and data rate are much higher than other communication links. Also Optical fibers are

*usman.illahi@ s.unikl.edu.my

very less susceptible to different type of losses, which are very common in copper cables; these losses include electromagnetic interferences and some other undesirable factors because of such advantages, optical fibers are used for Long Haul Communication [1].

The requirement for high capacity is root cause of important inventions done in 1990's i.e. optical amplifiers. EDFA (Erbium Doped Fiber Amplifier) is one of them and that different optical multiplexing techniques were invented as well like Wavelength Division Multiplexing (WDM). The WDM is the multiplexing technique that utilizes the bandwidth provided by the optical communication link by combining the multiple optical carries with different wavelength for each user. Nowadays WDM system is used by every optical communication network because of which the data rate is enhanced tremendously.

The capacity of wavelength division multiplexing (WDM) communication network can either be enhanced by increasing the number of channels or by the higher signal bit rate of time division multiplexing (TDM) [2]. In long haul high speed optical communication systems, the multiplexing techniques always play a significant role. Strong optical filtering is applied at the process of demodulation and only spectral efficient technique will perform well. As to counter the effect of distortion due to filtering, compact spectrum is required. The system performance and its efficiency depend on the multiplexing format used [3].

II. Literature Review

To enhance the capacity and efficiency of transmission by WDM different ideas were proposed. Among all these methods the advancement in the modulation technique is reported as the best solution to the problem [4, 5]. A lot of multi-level multiplexing techniques are introduced such as AM-Phase Shift Keying polybinary, polyquaternary and M-ary and they responded quite well against chromatic dispersion because of narrow spectral width [6], but there is a major disadvantage of multilevel coding techniques that they reduce the receiver sensitivity [7].

Pulse position modulation is another modulation format in the field of fiber optic communication. The logic behind this coding technique is to divide the symbol duration into small fragments. The symbols are sent during these small intervals by the transmitter. The amplitude and width of the Signal remains constant. Each pulse position is changed with respect to repeated pulse position. PPM required wider bandwidth and less optical power as compared to RZ or NRZ [8].

A new technique that can be used for optical communication called multi slot amplitude coding is introduced recently. This technique behaved significantly well at the time of clock recovery because initially zero level is used for all symbols. But the main limitation of this technique is the reduced receiver sensitivity [9]. The improvement in performance is reported by use of return-to-zero (RZ), but with reduced spectral efficiency [10,11].

But in common that modulation format is best for high speed long-haul communication links using WDM, which shows high spectral efficiency and better tolerance against dispersion [12]. So from previously discussed techniques it is concluded that there is still a need for new multiplexing format for fiber optic communication to handle all issues like better bandwidth utilization, complexity of the system, cost reduction and limitation of communication like dispersion tolerance, spectral efficiency and optical filtering.

III. Problem Statement

Currently, gain bandwidth of EDFA is limiting the maximum number of channels in WDM [12]. To enhance the capacity within the gain bandwidth limitations needs improvement in the spectral efficiency. To acquire the desired results optical filter with high stability along with narrow optical filtering are needed. A problem is associated with narrow optical filtering because it blocks the major part of the signal spectrum and that is distortion [13-18]. The solution to this problem, which is caused by narrow optical filtering, is high spectral efficiency achieved by reduced spectral width of the transmitted signal.

IV. Proposed Multiplexing Technique

A novel mapping multiplexing technique also called Tributary Mapping Multiplexing (TMM) can accommodate a lot of users per WDM channel [19]. Narrow optical filtering can be implemented by use of this technique ,which provides an opportunity to increase the number of channel within the allocated bandwidth to enhance the capacity. In this article a comparison of TMM and conventional TDM using RZ signaling with 50% duty cycle is made. The bit rate that is used per wavelength is 40 Gbit/s. For ease of discussion throughout the article conventional TDM using RZ with 50% duty cycle is referred to as TDM. Also the tolerance against chromatic dispersion is computed and discussed. We demonstrated that TMM perform much better in terms of spectral efficiency, dispersion tolerance and tolerance against narrow optical filtering.

V. TMM Unique Symbol Format

TMM is an efficient alternative technique for 40 Gbit/s optical transmission systems. TMM is a unique transmission technique that divides whole bit duration into two slots and different levels of amplitudes for data from different users over a single fiber. Number of users determines the number of amplitude levels [19].

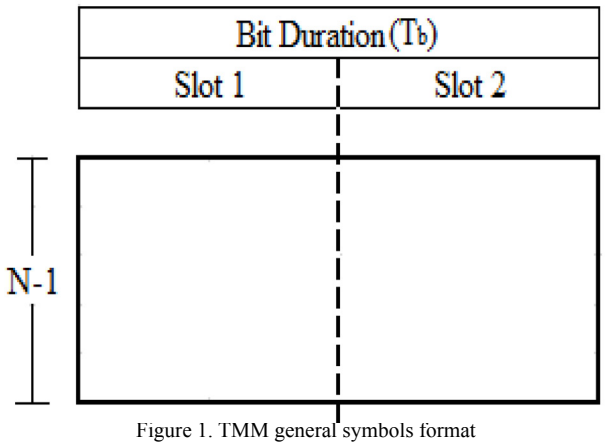


Figure 1. TMM general symbols format

Figure 1 is the representation of TMM general symbol format. Using TMM symbol formatting if N is the number of users then number of amplitude levels will be N-1 and two slots. Each slot duration is explained in [19].

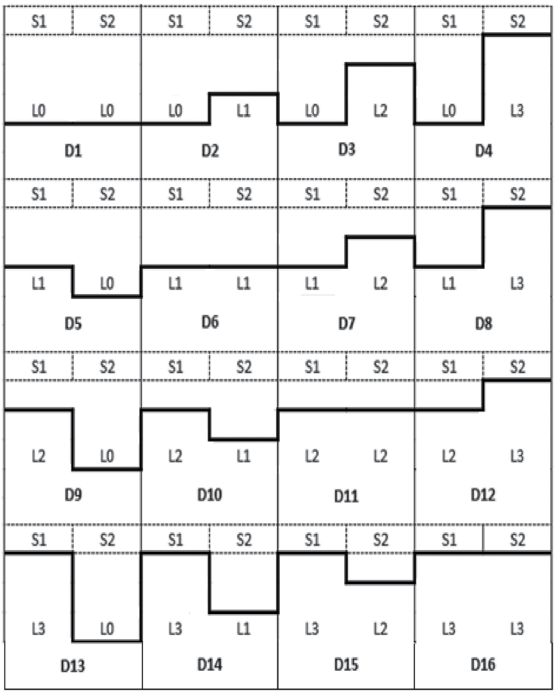
$$T_s = T_b / 2 \quad (1)$$

Where Tb represents the whole bit duration (1/bitrate). For sake of multiplexing different user by TMM technique a method known as mapping method is used. This mapping method coverts the data from different users into unique TMM symbols [19]

The 4-User Transmission system by using TMM as multiplexing method is represented by Figure 2. All possible combinations (D1 to D16) for 4-user multiplexed by TMM is shown is Figure 2 (a). The unique symbol of TMM is generated by mapping each of these possible combinations [19].

Case	U1	U2	U3	U4
D1	0	0	0	0
D2	0	0	0	1
D3	0	0	1	0
D4	0	0	1	1
D5	0	1	0	0
D6	0	1	0	1
D7	0	1	1	0
D8	0	1	1	1
D9	1	0	0	0
D10	1	0	0	1
D11	1	0	1	0
D12	1	0	1	1
D13	1	1	0	0
D14	1	1	0	1
D15	1	1	1	0
D16	1	1	1	1

(a)



(b)

Figure 2. (a) 16 possible combinations for 4-Users TMM System, (b) Unique TMM symbols associated to D1-D16

VI. Simulation Setup for Proposed Model of TMM System

In this research professional software tools, OptiSystem and MATLAB are used to evaluate the system performance. The evaluation of the system performance is used on Bit Error rate (BER), reported in [20].

A. Single Channel 40 Gbit/s TMM System

Figure 3 represents the implementation of 4-User TMM system over single wavelength channel. Each user is an RZ signal with 50% duty cycle and is transmitting at a rate of 10 Gbit/s. Data is generated by random bit sequence and in this way, U1, U2, U3 & U4 data is generated. This is done with the aid of Pseudo Random Bit Sequence (PRBS) of 210-1 along with RZ having 50% duty cycle.

As reported in [19] the multiplexer generates 16 unique symbols for four users by mapping algorithm. This algorithm is based on the combination of bits shown in Figure 2 (b).

At transmitting end, the generated symbols are modulated by Laser Diode (LD) at 1550 nm wavelength and Mach-Zehnder modulator (MZM).

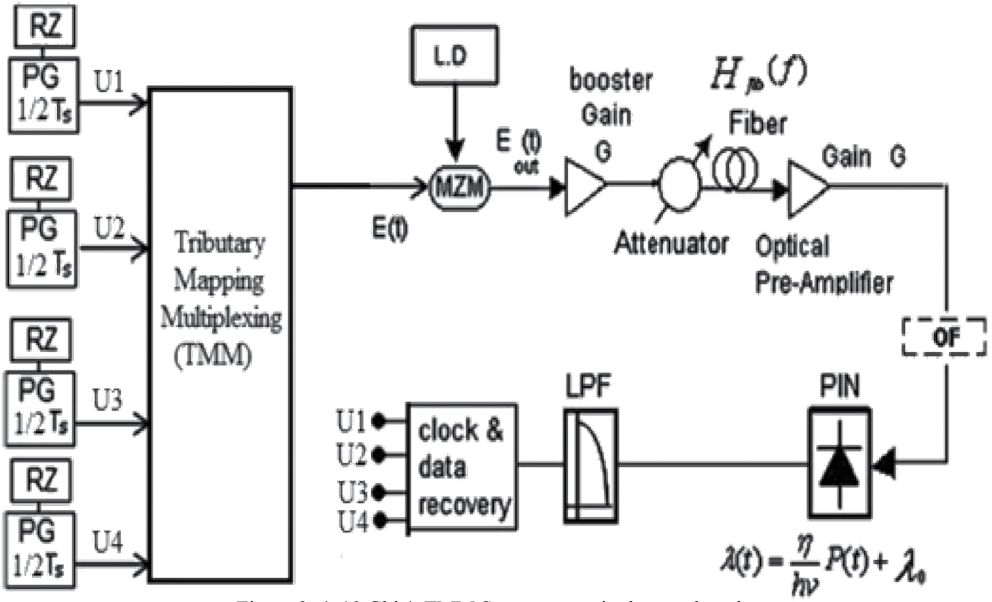


Figure 3. 4x10 Gbit/s TMM Setup over a single wavelength

Shown in Figure 3 a linear band pass system represented by a low-pass equivalent is used to modulate the fiber in which the dispersion of the fiber depends upon the nonlinear phase response of the transfer function of the fibers, reported in [21].

B. Data Recovery Rules

At receiving end photodiode is used to detect the optical signal. The received signal then passed through a low pass-filter (LPF) and then it is further passed from a Clock-and-Data-Recovery (CDR) unit to recover the received data [19]. Figure 4 represents the eye diagram of the output of the modulator.

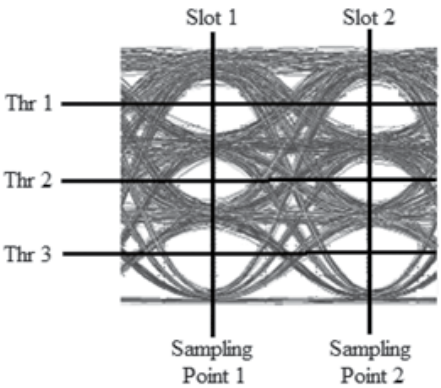


Figure 4. TMM Eye Diagram, including three thresholds and two sampling points

In CDR unit the detected signal is passed into the sampling circuit. In every symbol there are two slots, samples are taken at two sampling points S1 and S2 at each slot (Table 1). Sampling circuit outputs are fed in to the decision and regeneration unit. In this unit the sampled values are measured against three pre-defined threshold values, thr1, thr2 and thr3 (Table 1) and decisions are made on the basis of process defined in the Table 1:

TABLE I
TMM USER RECOGNITION RULES

Number	Rules			U1	U2	U3	U4
1	if S1 < thr1	then	S2 < thr1	0	0	0	0
			thr1 < S2 < thr2	0	0	0	1
			thr2 < S2 < thr3	0	0	1	0
			S2 > thr3	0	0	1	1
2	if thr1 < S1 <thr2	then	S2 < thr1	0	1	0	0
			thr1 < S2 < thr2	0	1	0	1
			thr2 < S2 < thr3	0	1	1	0
			S2 > thr3	0	1	1	1
3	if thr2 < S1 < thr3	then	S2 < thr1	1	0	0	0
			thr1 < S2 < thr2	1	0	0	1
			thr2 < S2 <thr3	1	0	1	0
			S2 > thr3	1	0	1	1
4	if S1 > thr3	then	S2 < thr1	1	1	0	0
			thr1 < S2 < thr2	1	1	0	1
			thr2 < S2 < thr3	1	1	1	0
			S2 > thr3	1	1	1	1

The original data by each user is reconstructed by data recovery unit, which uses the rules defined in table. For example, the regenerated data for U1 is binary 0 when the value taken at sampling point S1 is less than thr2 and sampling point S2 is greater or less than thr1, thr2 and thr3. Similarly, the regenerated data for U1 is binary 1 when the value taken at sampling point S1 is greater than thr2 and sampling point S2 is greater or less than thr1, thr2 and thr3. Similarly, data for U2, U3 and U4 can be regenerated [19].

VII. RESULTS AND DISCUSSIONS

A. Spectral Efficiency

Figure 5 shows the comparison between the optical spectra of TMM and conventional TDM at a bit rate of 40 Gbit/s. The spectral efficiency is an important tool that gives an idea about the capacity of digital communication systems. As represented clearly that main lobe of TMM signal is much narrower than conventional TDM signal and side lobes are compressed significantly. This is because TMM divides the whole bit duration into two equal slots and number of levels depending

on the number of users. Thus TMM requires a null-to-null spectral width of 2x [No. of Slots x single channel bit rate]/No. of Levels per slot, whereas, TDM requires 2 x [2 x aggregate bit rate]. The narrow spectral width of TMM gives a lot advantages when implemented on WDM system. These advantages are discussed in upcoming sections.

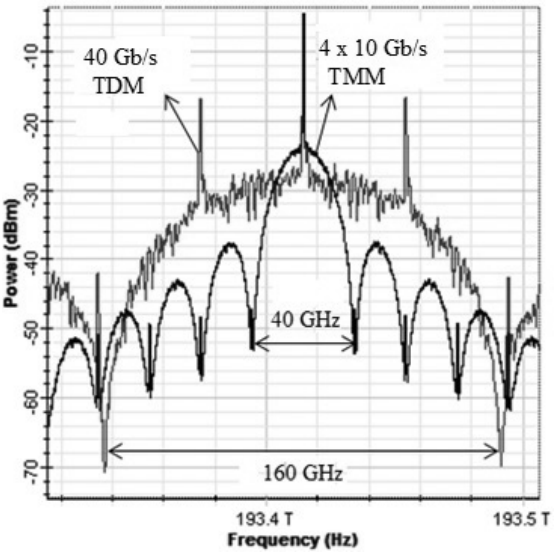


Figure 5. Optical Spectrum of 4 x 10 Gbit/s TMM vs 40 Gbit/s TDM

B. Tolerance against Narrow Optical Filtering

In the case of the communication systems working at an aggregate bit rate of 40 Gbit/s with broader optical spectrum the effect of optical filtering is an important issue. In the de multiplexing process of WDM strong optical filtering is applied. In this section the performance of TMM and TDM under strong filtering condition at 40 Gbit/s is observed.

The spectral width of TMM is narrow, because of which narrow band pass optical filter can be applied for de multiplexing process. Narrow filtering provide space and number of channels can be increased. But very narrow band pass optical filter is used, it blocks the major part of the signal spectrum and this causes the degradation of optical signal.

To compute the effect of distortion because of narrow filtering on optical signal, study measure the power penalties. The figure 6 shows the comparison between the TMM and TDM at 40 Gbit/s. In this figure power penalties are plotted against the optical filter bandwidth at a BER 10-9.

As shown by the Figure 6 the TMM is more robust to narrow optical filter as compared to conventional TDM. For 40 Gbit/s TMM system an optical filter with a bandwidth of 30 GHz can be used for 1 dB penalty. But for TDM optical filter bandwidth should be greater than 78 GHz.

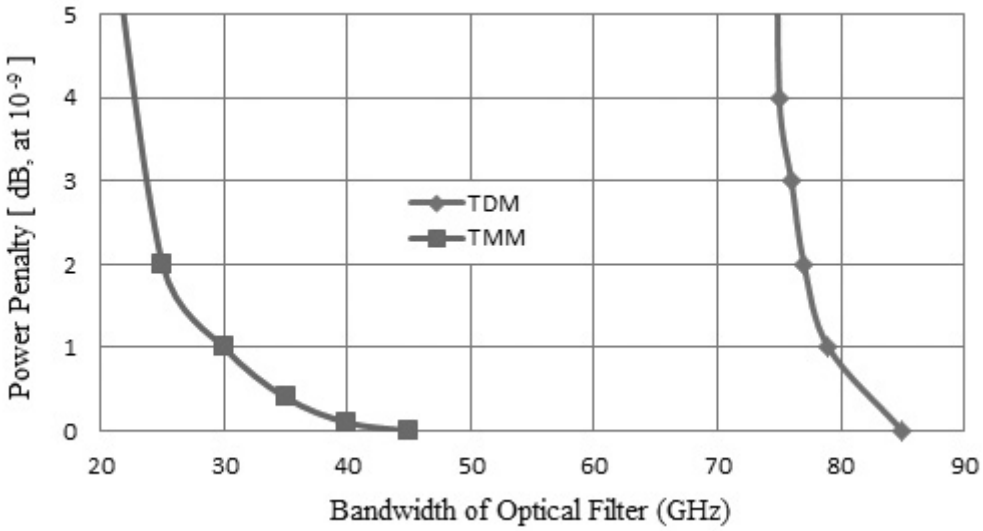


Figure 6. Tolerance against narrow optical filtering

C. Dispersion Tolerance

Here the comparison between the dispersion tolerances of TMM and conventional TDM is performed at aggregate bit rate of 40 Gbit/s. Figure 7 shows the positive and negative dispersion tolerances of TMM and TDM.

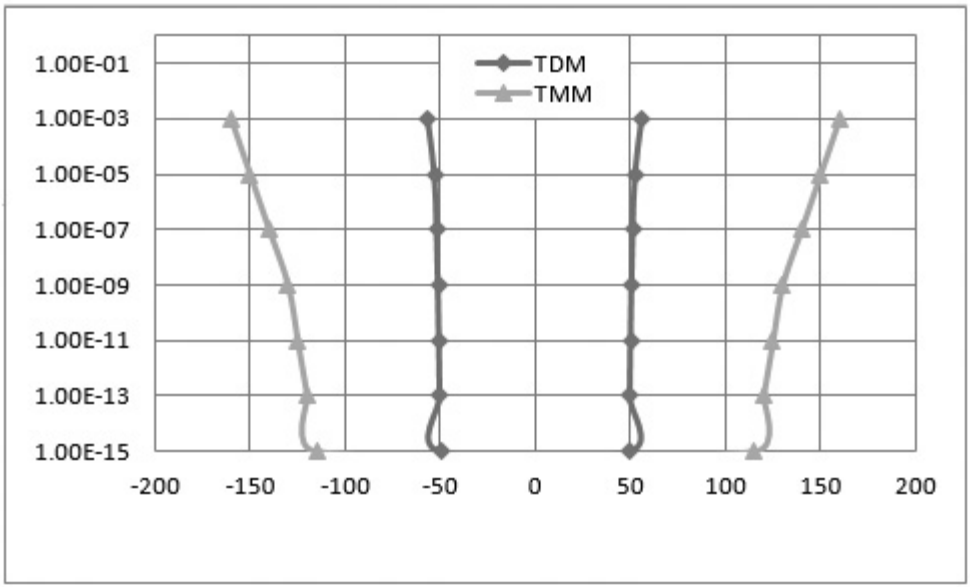


Figure 7. Chromatic dispersion tolerance of 4 x 10 Gbit/s TMM and 40 Gbit/s TDM at same received power

As demonstrated clearly that chromatic dispersion of TMM is ± 130 ps/nm and TDM is ± 51 ps/nm. So it is obvious that TMM is more robust to dispersion as compared to conventional TDM. The reason behind this tolerance is the smaller spectral width of TMM. The spectral width of TMM is 40 GHz, whereas the spectral width of TDM is 160 GHz. So if the spectral width will be smaller system will be more immune to Chromatic Dispersion.

VIII. Conclusion

The performance of 40 Gbit/s TMM and TDM over single channel WDM systems are demonstrated in this paper. The narrow optical spectrum of TMM provides tolerance against strong optical filtering applied at de-multiplexing process of WDM. The possibility of using an optical filter as narrow as 40 GHz for 40 Gbit/s TMM signal was confirmed. The high and symmetrical dispersion tolerance of ±130 ps/nm was obtained for TMM system, which is very attractive in high-speed WDM transmission systems. Based on the simulation results TMM system demonstrated a better spectral width and tolerance against optical filtering in comparison with conventional TDM at 40 Gbit/s bitrate. So, it is concluded that by using TMM signal format the performance and capacity of WDM transmission system can be enhanced significantly.

Acknowledgments

The authors would like to acknowledge the Postgraduate and Research Department of UniKL BMI for providing the research facilities and Gomal University, Pakistan for granting study leave to carry out this research.

References

[1] Zheng, Z., An, L., Li, Z., Zhao, X., & Liu, X. (2008). All-optical regeneration of DQPSK/QPSK signals based on phase-sensitive amplification. Optics Communications, 281(10), 2755-2759.W.-K. Chen, Linear Networks and Systems (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[2] Chow, C. W., & Yeh, C. H. (2009). Signal remodulation without power sacrifice for carrier distributed hybrid WDM-TDM PONs using PolSK. Optics Communications, 282(7), 1294-1297.

[3] A. Malekmohammadi, G. A. Mahdiraji , A. F.Abas, M. K. Abdullah, M. Mokhtar, M. F. A. Rasid, "Realization of High Capacity Transmission in Fiber Optic Communication Systems Using Absolute Polar Duty Cycle Division Multiplexing (AP-DCDM) Technique," Optical Fiber Technology, vol.15, 2008.

[4] Winzer, P. J., & Essiambre, R. J. (2006). Advanced modulation formats for high-capacity optical transport networks. Journal of Lightwave Technology, 24(12), 4711-4728.

[5] Shao, Y., Chen, L., Wen, S., Xiao, Y., Cheng, L., Xu, H., & Pi, Y. (2008). Novel optical orthogonally modulation scheme for superimposing DPSK signals on dark RZ signals. Optics Communications, 281(14), 3658-3667.

[6] Malekmohammadi, A., Abas, A. F., Abdullah, M. K., Mahdiraji, G. A., Mokhtar, M., & Rasid, M. F. A. (2009). Absolute polar duty cycle division multiplexing over wavelength division multiplexing system. Optics Communications,

282(21), 4233-4241.

[7] Ohm, M., & Speidel, J. (2003). Quaternary optical ASK-DPSK and receivers with direct detection. IEEE Photonics Technology Letters, 15(1), 159-161.

[8] Da Silva, J. I., & Sombra, A. S. B. (1998). Pulse position modulation (PPM) of ultrashort pulse trains in optical fibers. Optics communications, 152(1), 59-64.

[9] Talib, R., Abdullah, M. F. L., Malekmohammadi, A., & Abdullah, M. K. (2011, July). Multi-slot and multi-level coding technique over amplitude-shift keying modulation for optical communication links. In Networks and Optical Communications (NOC), 2011 16th European Conference on (pp. 161-164). IEEE.

[10] Yang, W., Zhang, M., & Ye, P. (2009). Analysis of 160Gb/s all-optical NRZ-to-RZ data format conversion using quantum-dot semiconductor optical amplifiers assisted Mach-Zehnder interferometer. Optics Communications, 282(9), 1744-1750.

[11] Breuer, D., & Petermann, K. (1997). Comparison of NRZ-and RZ-modulation format for 40-Gb/s TDM standard-fiber systems. IEEE Photonics Technology Letters, 9(3), 398-400.

[12] Ono, T., Yano, Y., Fukuchi, K., Ito, T., Yamazaki, H., Yamaguchi, M., & Emura, K. (1998). Characteristics of optical duobinary signals in terabit/s capacity, high-spectral efficiency WDM systems. Journal of Lightwave Technology, 16(5), 788.

[13] Li, M., Fujii, T., Li, H., & Painchaud, Y. (2009). Proposal and realization for a broadband all-fiber non-uniformly spaced multi-channel optical filter. Optics Communications, 282(5), 879-882.

[14] A Muhammad, "Power Management of Portable Devices by Using Clutter Based Information" IJCSNS, International Journal of Computer Science and Network Security, VOL. 9 No. 4, April 2009, tel-00815919, version 1-19 Apr 2013 pp 237-244.

[15] A Muhammad, MS Mazliham, M Shahrulniza, M Amir, "Posterior Probabilities based Location Estimation (P2LE) Algorithm for Locating a Mobile Node in a Disaster Area M" MULTICONF p 13-16, 2009.

[16] Muhammad Alam, Mazliham Muhammad Suud, Patrice Boursier, Shahrulniza Musa, Jawahir Che Mustapha Yusuf, "Predicted and corrected location estimation of mobile nodes based on the combination of kalman filter and the bayesian decision theory" , International Conference on Mobile Wireless Middleware, Operating Systems, and Applications, Pages 313-325, Publisher Springer Berlin Heidelberg

[17] A Muhammad, MS Mazliham, Patrice Boursier, M Shahrulniza, Jawahir Che Mustapha Yusuf, "Terrain/clutter based location prediction by using multi-condition Bayesian decision theory" Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication. P. 110-117, 2012.

[18] Muhammad Alam, "Error corrected location determination in an outdoor wireless environment by using estimation, filtering, prediction and fusion techniques : A Wi-Fi application by using terrain based knowledge" , HAL, 2013 pages 313.

[19] Malek mohammadi, A., Illahi, U., Saqlain, M., & Abdullah, M. K. (2012, October). A novel mapping multiplexing technique for high speed optical fiber communication systems. In 2012 IEEE 3rd International Conference on Photonics (pp. 323-326). IEEE.

[20] Saqlain, M., Illahi, U., Iqbal, J., & Ali, E. (2013). Comprehensive Study Analysis of Novel Mapping Multiplexing (NMM) Technique for Long Haul Optical Fiber Transmission Systems. European Scientific Journal, 9(30).

[21] Avlonitis, N., Yeatman, E. M., Jones, M., & Hadjifotiou, A. (2006). Multilevel amplitude shift keying in dispersion uncompensated optical systems. IEE Proceedings-Optoelectronics, 153(3), 101-108.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Ranking and Sensitivity Analysis of Key Factors for Successful Project Management Performance: An Application of AHP for Oil and Gas Sector

*Farhaj Ishtiaq and Mirza Jahanzaib
Department of Engineering Management, U.E.T Taxila, Pakistan

***Abstract*- Complexities faced by oil and gas projects due to uncertainty and risk, demand the implementation of project management techniques for their successful completion. Therefore, this is made by using analytical hierarchy process, to identify and prioritize the key factors for successful project management performance of oil and gas projects. These factors are categorized into three groups which include attributes of project staff, project planning process and assessment of project quality. Using expert choice, a hierarchy is developed followed by pairwise comparison based upon data collection from industrial experts of oil and gas sector. Results of analytical hierarchy process (AHP) concluded that, project completion within estimated time and budget, clarity of objectives and involvement of top management are most crucial elements for improvement in project management performance of oil and gas projects. Whereas sensitivity analysis being carried out according to three different scenarios highlighted factors according to their relative importance.**

Keywords: Oil and Gas sector, Project Management, Analytical Hierarchy Process and Sensitivity Analysis.

I. Introduction

Oil and gas sector is considered as major contributor of nation's economy and infrastructural development [1]. This sector has two major divisions i.e. upstream and downstream. Upstream sector is concerned with exploration and production of oil and gas and downstream sector deals with refining, transportation and marketing [2]. Oil and gas sector is known by certain characteristics such as huge investments, environmental effects, multi discipline workforce, global influence and high rewards. Many times oil and gas projects face complexities due to unstable political situation, increased market demand, fluctuations in price and tough schedule [3]. These issues sometimes lead oil and gas projects towards cost and schedule overrun as well. The reasons behind cost and schedule overrun are unavailability of skilled staff, unclear definition of projects scope, inappropriate planning, poor project control, lack of competent leadership and inexperienced project management personnel [4]. Besides this, these projects face issues like lack of trained staff, unfavorable market

*farhajch619@yahoo.com

conditions and environmental concerns. Project of oil and gas industry are usually risky, uncertain and provide intangible benefits. These risk also exists due to uncertain cash flow and irreversibility of these projects incorporating economic risks as well [5].

Project management tools and techniques are highly important to handle complex projects within estimated time and budget [6]. It is used by organizations to handle frequent customer needs with in allocated timeframes along with fast decision making. A study based in UAE has concluded that project management practices have positive influence on project success [7]. Organizations use project management tools and techniques to achieve organizational goals in a focused manner. It is also used to predict crisis while handling uncertainties to make a project successful [8]. The success of project management process is analyzed by defined criteria based on cost, time and quality, whereas project success is measured by its objectives [9].

A project which is executed in right direction has an ability to be successful but successful project management always enhance success of projects [10]. Project management practices not only enhance the performance of project manager who is using it, but it also improves project performance. It improves project performance by proper budget control and time utilization as estimated [11]. Project management tools and techniques helps to complete projects with in estimated budget and time while meeting desired quality level [12]. Its tools and techniques (DMAIC, PDCA, risk map, decision tree, sensitivity analysis, SWOT analysis, cause and effect diagram) are also used by oil and gas sectors for successful completion of projects [13].

According to the literature, AHP is widely used for many purposes relevant to project and project management. It is one of the most promising technique used for multi criteria decision making [14]. In a study AHP has been used for the selection of investment project of solar thermal based power plant [15]. The application of AHP also exists for construction projects to access safety risk during planning and budgeting phase [16]. A study carried out in Italy used AHP for the assessment of hydropower projects by incorporating stakeholders [17]. AHP also has its application for identification and evaluation of critical success factors for projects of construction industry [18]. The performance measurement of green supply chain of a manufacturing organization is also made using AHP [19]. In table 1, studies applying AHP for different projects is given.

TABLE 1
STUDIES USING AHP FOR DIFFERENT PROJECTS

Reference	Topic Addressed
[15]	Selection of investment project of solar thermal based power plant
[16]	Assessment of safety risk during planning and budgeting phase of construction project
[17]	Assessment of hydropower projects by incorporating stakeholders
[18]	Identification and evaluation of critical success factors for projects of construction industry
[19]	Performance measurement of green supply chain of a manufacturing organization
[21]	Evaluation of complexity of projects
[22]	Selection of a renewable energy project in Spain
[23]	Management of project risk for construction projects in India
[24]	Project selection process for six sigma deployment
[25]	Risk assessment for construction projects in China

Project management performance is directly associated with success of oil and gas projects. The knowledge of factors essential for improvement in project management process is helpful for project managers for execution and successful completion of oil and gas projects. Therefore, keeping in view, issues faced by oil and gas sector, this study is made using AHP to identify and prioritize the factors, which are essential for the successful project management performance of oil and gas projects.

II. ANALYTICAL HIERARCHY PROCESS

AHP is a multi-criteria decision making process which helps organizations to deal with complex and multiple conflicting objectives. It is widely used in many fields like engineering, manufacturing, management and social sciences. It uses pairwise comparison to rank alternatives subject to particular goals [20]. Analytical hierarchical process is a systematic way to prioritize and weight all the objectives. It is assumed that all the objectives of a particular problem are represented in a hierarchy. This technique has ability to deal with complex phenomena of real life by producing most consistent results. Analytical hierarchical process also has a potential for linking with linear programming and expert's systems. It also facilitates decision makers to tradeoff between criteria. It has following basic steps.

1) A hierarchy based structure is defined for identified problem by decomposing it into goal, criteria and sub criteria. It is most important and fundamental step of decision making process. Basically hierarchy based structure is used to link elements of one level to next associated level.

2) After the development of hierarchy, pair wise comparison is made between all the

alternatives by expert's / decision makers. This comparison is made based on a scale, according to which decision maker's rate elements [14, 26]. Description of scale is given in table 2.

3) Pairwise comparisons of previous step are synthesized to get result of overall priorities and weights of elements with respect to the goals.

AHP is developed using a software known as "Expert Choice". It allows group decision making to solve complex phenomena by sharing experience and knowledge. There are certain benefits of Expert Choice, which are explained below.

I. It helps to minimize the influence of dominant group member or groupthink.

II. Overall structure of hierarchy is based upon agreement of whole group by considering their concerns. With group discussion, modifications can be made to cover all the aspects.

III. In a situation where it becomes difficult to reach a conclusion, it may be decided through voting or average of judgments may be taken.

IV. It synthesizes the objectives with respect to goal to get overall priorities.

V. Sensitivity analysis is performed using Expert Choice to observe the result of change in objectives.

VI. It is an ideal tool for group decisions through cohesive and rigorous process.

TABLE 2
SCALE FOR PAIRWISE COMPARISON BETWEEN FACTORS

Level of Importance	**Definition**	**Interpretation**
1	Equally preferred	Two activities contribute equally to the objective
3	Moderately	Experience and judgment slightly favor one activity over another
5	Strongly	Experience and judgment strongly or essentially favor one activity over another
7	Very strongly	An activity is strongly favored over another and its dominance demonstrated in practice
9	Extremely	The evidence favoring one activity over another is of the highest degree possible for affirmation
2,4,6,8	Intermediate values	Used to represent a compromise between preferences listed above
Reciprocals	Reciprocals for inverse comparison	

III. METHODOLOGY

Key factors for successful project performance are identified with the help of literature review and discussion with experts of oil and gas sector. After which a hierarchy is developed based upon three level. Data collection is made by pair wise comparison of one factor with other factor according to their relative importance using scale ranging from 0 to 9. Then based on this data collection, key

factors are prioritized according to their importance. After which sensitivity analysis is carried out considering different scenarios to help project managers to deal with varying conditions of oil and gas projects.

A. Key Factors of Successful Project Management Performance

There are many factors which influence project management performance to various extent. These indicators are identified with the help of literature and expert opinion. In this study, these factors are grouped into three categories which include attributes of project staff, project planning process and assessment of project quality.

Several studies indicates that technical knowledge, collaboration between technical and non-technical staff, training of staff and leaderships skills of team are important indicators of workforce's attributes [1, 2, 27, 28]. Clarity of objectives, project completion with in estimated time and budget, work norms and standards and involvement of top management are most important factors for project planning [2, 27-30]. Several investigations concluded that project quality assessment is based upon implementation of quality control programs, ability to respond quickly and adequate risk analysis [2, 27, 30-32]. All these variables are also given in table 3.

TABLE 3
ESSENTIAL FACTOR FOR IMPROVEMENT IN PROJECT MANAGEMENT PERFORMANCE

Main	Sub Categories	Notation	References
Attributes of project staff (PS)	Technical knowledge Collaboration Leadership skills Training	PS1 PS2 PS3 PS4	[1, 2, 27, 28]
Project planning process (PP)	Project completion with in estimated time and budget Work norms and standards Clarity of objectives Top management involvement	PP1 PP2 PP3 PP4	[2, 27-30]
Assessment of project quality (PQ)	Quality control programs Ability to respond quickly Risk analysis Implementation of ISO standards	PQ1 PQ2 PQ3 PQ4	[2, 27, 30-32]

B. Development of AHP Model

Hierarchy model being developed by AHP must meet the goal behind it. The model developed for this study is based on three levels as shown in figure 1. First level is representing the goal of designed hierarchy i.e. key variables for successful project management performance. Objectives / criteria for achieving goal are represented by second level of hierarchy i.e. project staff, project planning process, project quality measures. Whereas sub criteria for objectives are defined at level three. Sub criteria for project staff include staff's expertise, collaboration, training and leadership

skills, whereas project planning process is sub categorized into project completion with in estimated time and budget, involvement of top management, project objectives clarity and its norms and standards. Sub criteria for project quality measures include quality control programs, quick response of queries, implementation of ISO standards and risk analysis.

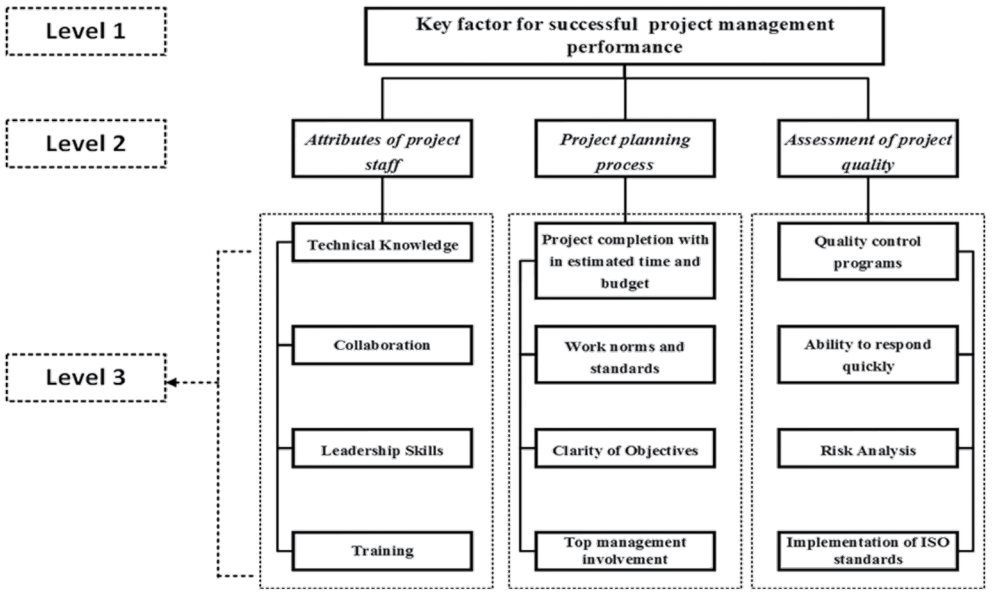


Figure 1: Hierarchy developed for essential factors of project management performance

C. Data Collection

A questionnaire based survey is carried out by industrial experts of oil and gas sector to determine the relative importance of each factor with respect to other. This questionnaire is based upon pair wise comparison between factors using a scale of 0 to 9. Each element is compared with other element to evaluate its relative importance based on some goal / criteria [33]. This method also allows cross checking and consistency between elements. First, a pairwise comparison is made between elements of second level according to the goal of this study. Attributes of project team, project planning process and assessment of project quality are compared with each other. After which, at third level three pairwise comparisons are made for each element of first level according to their respective sub criteria. Pairwise comparison between factors of first level is given in table 4. Whereas comparisons between attributes of project staff, project planning process, and assessment of project quality are given in table 5, 6 and 7 respectively.

TABLE 4
PAIRWISE COMPARISON BETWEEN FACTORS OF FIRST LEVEL

	PS	PP	QP	Priorities
PS		0.5	2	0.327
PP			1	0.413
QP				0.260

TABLE 5
PAIRWISE COMPARISON BETWEEN SUB CATEGORIES OF "ATTRIBUTES OF PROJECT STAFF"

	PS1	PS2	PS3	PS4	Priorities
PS1		0.67	1.47	1.58	0.281
PS2			0.70	0.86	0.249
PS3				1.18	0.254
PS4					0.216

TABLE 6
PAIRWISE COMPARISON BETWEEN SUB CATEGORIES OF "PROJECT PLANNING PROCESS"

	PP1	PP2	PP3	PP4	Priorities
PP1		1.29	1.33	1.58	0.286
PP2			0.69	0.97	0.204
PP3				.90	0.277
PP4					0.233

TABLE 7
PAIRWISE COMPARISON BETWEEN SUB CATEGORIES OF "ASSESSMENT OF PROJECT QUALITY"

	QP1	QP2	QP3	QP4	Priorities
QP1		1.09	2.31	1.59	0.344
QP2			1.40	1.37	0.278
QP3				0.97	0.179
QP4					0.200

D. Ranking of Key Factors using AHP

Pairwise comparison between factors is synthesized to get overall ranking of variables. For attributes of project staff, technical expertise of workforce is the most important element followed by leadership skills, collaboration and training respectively as shown in figure 2. Whereas for project planning process, the most important attribute is project completion within estimated budget and time followed by clarity of objectives, involvement of top management and work norms and standards respectively as indicated by figure 3. Assessment of project quality is based upon quality control programs, quick response of queries, implementation of ISO standards and risk analysis respectively as shown in figure 4. Overall consistency of all measures is less than cut off value of 0.20 [34].

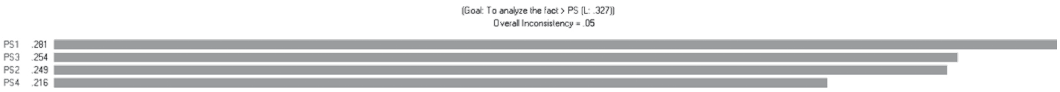


Figure 2: Prioritization of factors of "attributes of project staff"

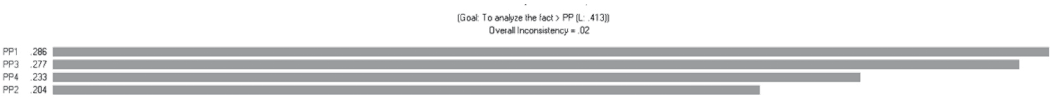


Figure 3: Prioritization of factors of "project planning process"

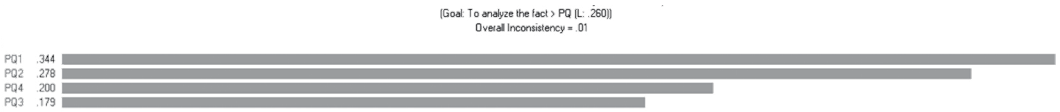


Figure 4: Prioritization of factors of "assessment of project quality"

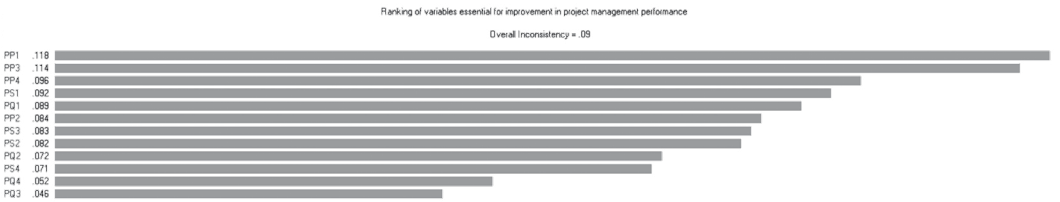


Figure 5: Overall Prioritization of factors for successful project management performance

According to the results produced by analytical hierarchical process, most essential element for successful performance of project management process is project completion within expected time and budget followed by clarity of objectives and involvement of top management. Whereas least significant contributors for successful project management process are training of staff, implementation of ISO standards and risk analysis respectively. In figure 5, all the factors contributing for better project management performance are shown in a sequence with consistency less than 0.1.

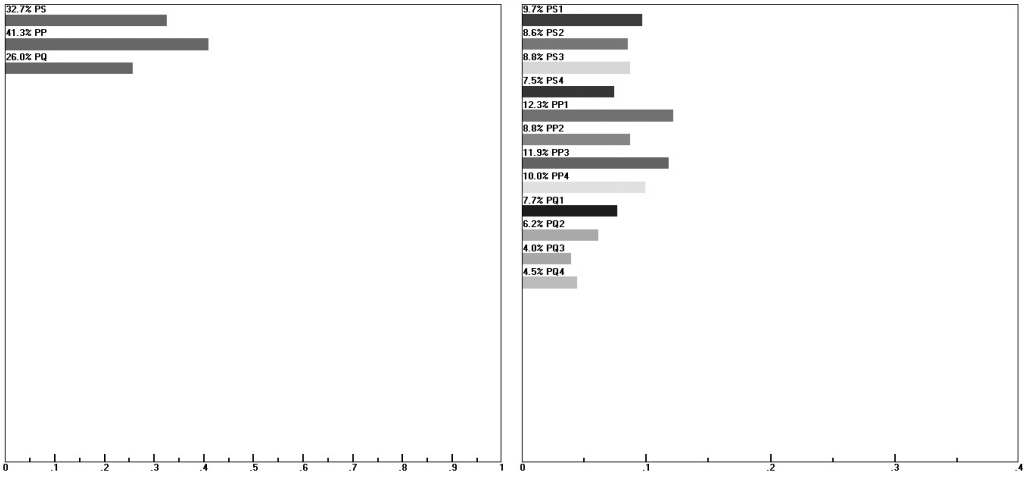
All variables are ranked according to their priority level in table 8.

TABLE 8
RANKING OF ALL FACTORS ESSENTIAL FOR PROJECT MANAGEMENT PERFORMANCE

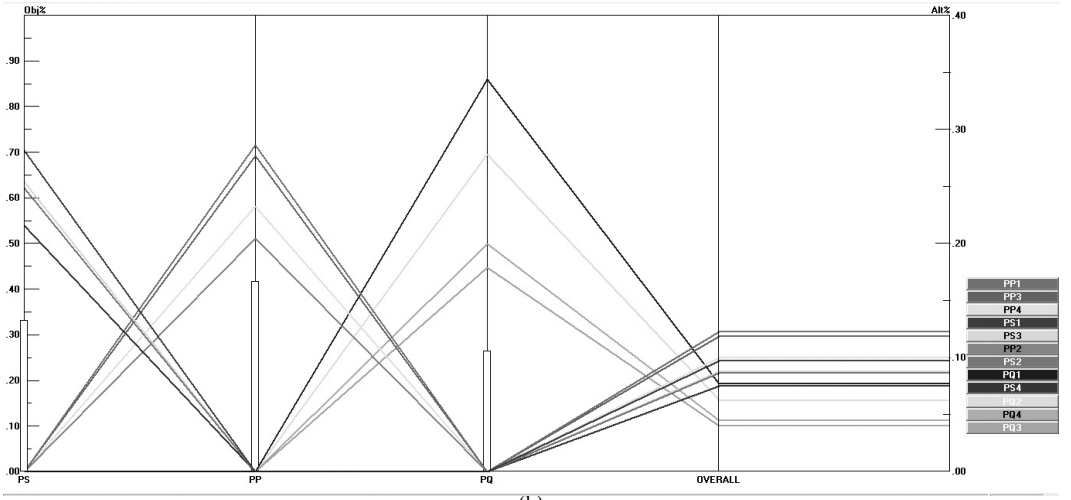
Main Factors	Sub Categories	Notation	Rank
Attributes of project staff (PS)	Technical knowledge	PS1	4
	Collaboration	PS2	8
	Leadership skills	PS3	7
	Training	PS4	10
Project planning process (PP)	Project completion with in estimated time and budget	PP1	1
	Work norms and standards	PP2	6
	Clarity of objectives	PP3	2
	Top management involvement	PP4	3
Assessment of project quality (PQ)	Quality control programs	PQ1	5
	Ability to respond quickly	PQ2	9
	Risk analysis	PQ3	12
	Implementation of ISO standards	PQ4	11

E. Sensitivity Analysis

The last step of AHP based decision making is sensitivity, where input data is slightly changed to observe effect on overall results [35]. It is best performed with graphical interface to help decision makers. Sensitivity analysis of AHP based model is shown in figure 6(a&b) with overall ranking of all elements. In figure 6a, percentage contribution of each main factor and sub categories is also shown. Besides this, three additional scenarios are discussed by rearranging overall priority structure of model. Analysis of model with different priorities structure, helps experts to evaluate different policies before making a final decision.



(a)



(b)

Figure 6 (a & b): Sensitivity analysis of AHP Model

I. Sensitivity Analysis w.r.t "attributes of project staff"

For first scenario, "attributes of project team" is given highest priority followed by planning process and quality assessment as shown in figure 7. For this scenario, technical expertise of team is ranked at one followed leadership skills, collaboration and training respectively. Whereas least contributor for this scenario are quick response of queries, implementation of ISO standards and risk analysis. When factor "attributes of project staff" is dragged down by giving priority to other two factors, it is observed that project completion within estimated time and budget and clarity of objectives become dominant factors.

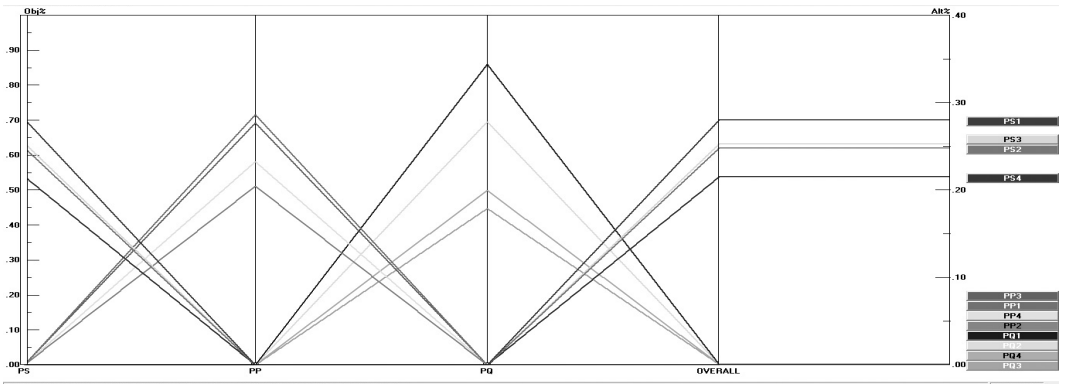


Figure 7: Sensitivity Analysis with priority to "attributes of project team"

II. Sensitivity Analysis w.r.t "project planning process"

In this scenario, project planning process is given highest priority followed by attributes of project staff and assessment of project quality as indicated in figure 8. According to this scenario, most significant contributor for project management performance are project completion within expected budget and time followed by clarity of objectives, top management involvement and work norms and standards. The least contributor for this scenario are training of team, quick response of queries, implementation of ISO standards and risk analysis. When factor "project planning process" is dragged down by giving priority to other two factors, it is concluded that staff's knowledge and leadership skills become dominant factors.

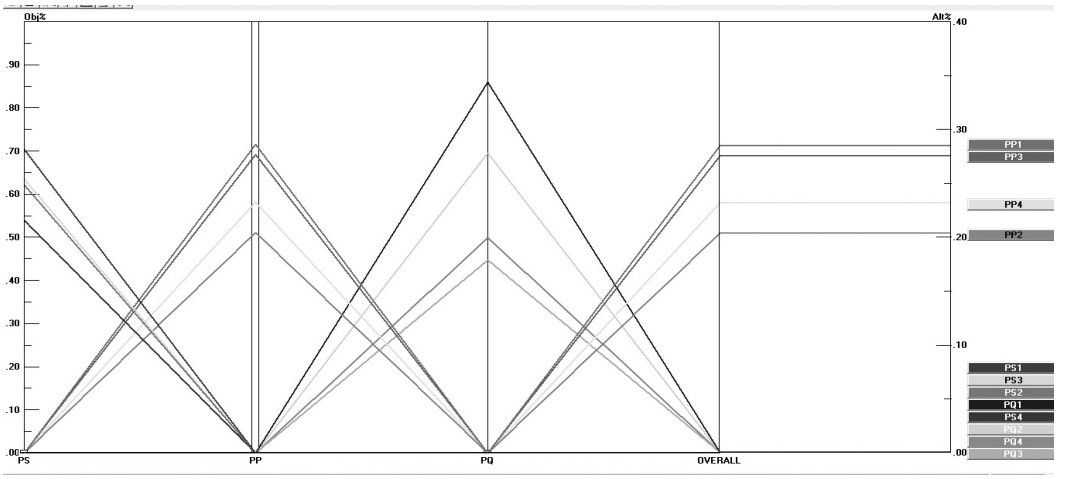


Figure 8: Sensitivity Analysis with priority to "project planning process"

III. Sensitivity Analysis w.r.t "assessment of project quality"

For third scenario, assessment of project quality is given highest priority followed by attributes of staff and project planning process as shown in figure 9. Most important factors for this scenario are quality control programs, quick response of queries, implementation of ISO standards and risk analysis respectively. Whereas least significant contributor are work norms and standards, leadership skills, collaboration and training of project staff. Whereas, when assessment of project quality is less prioritized with respect to other two factors then project completion within forecasted time and budget and technical knowledge of team becomes most dominant factors.

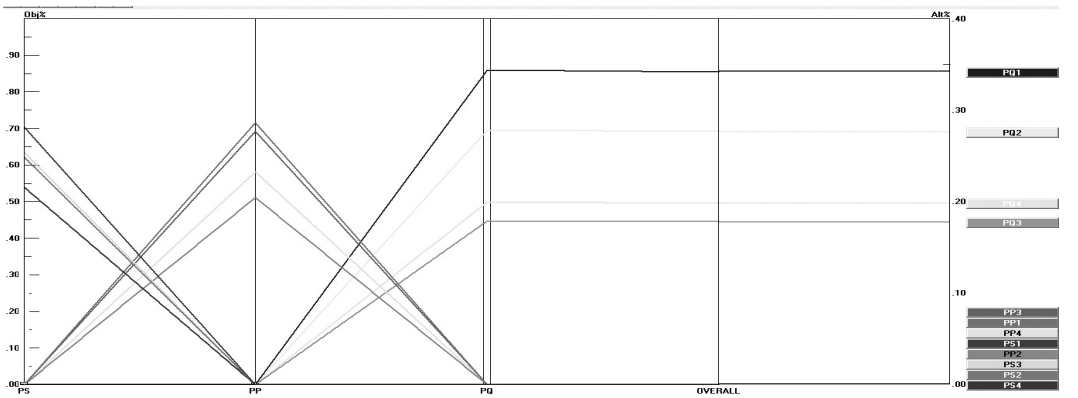


Figure 9: Sensitivity Analysis with priority to "assessment of project quality"

IV. Conclusion and Recommendations

Oil and gas project face many difficulties due to inappropriate planning, tight schedule and uncertainties. Therefore, this study is made using AHP to help project managers by ranking the key factors for successful project management performance. With the help of literature review and expert's opinion, key factors are selected. After which an analytical hierarchy based model is developed using these factors to facilitate oil and gas industrial experts for decision making in different scenarios. Data collection is made from oil and gas experts using a scale ranging from zero to nine. Based upon pair wise comparison on Expert Choice, collected data is synthesized to get overall results of hierarchy. After which it is found that

1) Project completion within expected time and budget followed by clarity of objectives and involvement of top management are most crucial elements for better project management performance of Oil and Gas projects.

2) Least significant factors for improvement in project management process are quick response of queries, implementation of ISO standards and risk analysis.

3) Three different scenarios are also analyzed in this study by sensitivity analysis to help project managers in varying conditions. Each scenario has different dominant and least contributing factors.

Therefore, project managers should focus on highlighted factors of this study, to achieve success for oil and gas project management process while handling all uncertainties. It will help project managers to minimize difficulties faced during execution of oil and gas projects. Data collection for this study is made from oil and gas sector of Pakistan, whereas for more generic results, data collection can also be made from oil and gas companies of other countries as well.

References

[1] Weijermars, R., Accelerating the three dimensions of E&P clockspeed–a novel strategy for optimizing utility in the oil & gas industry. *Applied Energy,* 2009. 86(10): p. 2222-2243.

[2] Salazar-Aramayo, J.L., et al., A conceptual model for project management of exploration and production in the oil and gas industry: The case of a Brazilian company. *International Journal of Project Management,* 2013. 31(4): p. 589-601.

[3] Noreng, O., Crude power: politics and the oil market. Vol. 21. 2006: IB Tauris.

[4] Jergeas, G.F. and J. Ruwanpura, Why cost and schedule overruns on mega oil sands projects? *Practice Periodical on Structural Design and Construction,* 2009. 15(1): p. 40-43.

[5] Wagner, J. and K. Armstrong, Managing environmental and social risks in international oil and gas projects: Perspectives on compliance. T*he Journal of World Energy Law & Business*, 2010. 3(2): p. 140-165.

[6] Madden, P.B. and J.D. Morawski, The future of the canadian oil stands: engineering and project management advances. *Energy & Environment*, 2011. 22(5): p. 579-596.

[7] Mir, F.A. and A.H. Pinnington, Exploring the value of project management: linking project management performance and project success. *International journal of project management*, 2014. 32(2): p. 202-217.

[8] Vrchota, J. and P. Řehoř, Project Management and the Importance of Crises in the Sectors of the National Economy. *Procedia Computer Science,* 2016. 100: p. 362-368.

[9] Joslin, R. and R. Müller, Relationships between a project management methodology and project success in different project governance contexts. *International Journal of Project Management,* 2015. 33(6): p. 1377-1392.

[10] Munns, A. and B.F. Bjeirmi, The role of project management in achieving project success. *International journal of project management,* 1996. 14(2): p. 81-87.

[11] Raymond, L. and F. Bergeron, Project management information systems: An empirical study of their impact on project managers and project success. *International Journal of Project Management*, 2008. 26(2): p. 213-220.

[12] Keri, E., S. Pearlson, and C. Saunders, *Managing and using information systems: A strategic approach*. 2006, USA: Wiley.

[13] Badiru, A.B. and S.O. Osisanya, *Project management for the oil and gas industry: a world system approach. 2016: CRC Press.*

[14] Zakria, G., et al., Selecting and prioritizing key factors for CAD/CAM software in small-and medium-sized enterprises using AHP. *Frontiers of Mechanical Engineering in China*, 2010. 5(2): p. 157-164.

[15] Aragonés-Beltrán, P., et al., An AHP (Analytic Hierarchy Process)/ANP (Analytic Network Process)-based multi-criteria decision approach for the selection of solar-thermal power plant investment projects. *Energy*, 2014. 66: p. 222-238.

[16] Aminbakhsh, S., M. Gunduz, and R. Sonmez, Safety risk assessment using analytic hierarchy process (AHP) during planning and budgeting of construction projects. *Journal of safety research,* 2013. 46: p. 99-105.

[17] Rosso, M., et al., Integrating multicriteria evaluation and stakeholders analysis for assessing hydropower projects. *Energy Policy*, 2014. 67: p. 870-881.

[18] Gudienė, N., et al., Identification and evaluation of the critical success factors for construction projects in Lithuania: AHP approach. *Journal of Civil Engineering and Management*, 2014. 20(3): p. 350-359.

[19] Dey, P.K. and W. Cheffi, Green supply chain performance measurement using the analytic hierarchy process: a comparative analysis of manufacturing organisations. *Production Planning & Control*, 2013. 24(8-9): p. 702-720.

[20] Masood, S.A., M. Jahanzaib, and K. Akhtar, Key Performance Indicators Prioritization in Whole Business Process: A Case of Manufacturing Industry. *Life Science Journal*, 2013. 10(4s): p. 195-201.

[21] Vidal, L.-A., F. Marle, and J.-C. Bocquet, Using a Delphi process and the Analytic Hierarchy Process (AHP) to evaluate the complexity of projects. *Expert systems with applications*, 2011. 38(5): p. 5388-5405.

[22] San Cristóbal, J., Multi-criteria decision-making in the selection of a renewable energy project in spain: The Vikor method. *Renewable energy*, 2011. 36(2): p. 498-502.

[23] Dey, P.K., Managing project risk using combined analytic hierarchy process and risk map. *Applied Soft Computing,* 2010. 10(4): p. 990-1000.

[24] Kumar, M., J. Antony, and B. Rae Cho, Project selection and its impact on the successful deployment of Six Sigma.

Business Process Management Journal, 2009. 15(5): p. 669-686.

[25] Zhang, G. and P.X. Zou, Fuzzy analytical hierarchy process risk assessment approach for joint venture construction projects in China. J*ournal of Construction Engineering and Management*, 2007. 133(10): p. 771-779.

[26] Mirza Jahanzaib, D. and K. Akhtar, Prioritizing Key Quality Indicators for Engineering Education using Analytical Hierarchical Process (AHP): Case Study.

[27] Doloi, H., K. Iyer, and A. Sawhney, Structural equation model for assessing impacts of contractor's performance on project success. *International Journal of Project Management*, 2011. 29(6): p. 687-695.

[28] Fortune, J. and D. White, Framing of project critical success factors by a systems model. *International journal of project management*, 2006. 24(1): p. 53-65.

[29] Xia, B., et al., Investigating the impact of project definition clarity on project performance: Structural equation modeling study. *Journal of Management in Engineering*, 2015. 32(1): p. 04015022.

[30] Xiong, B., et al., Examining the influence of participant performance factors on contractor satisfaction: A structural equation model. *International Journal of Project Management*, 2014. 32(3): p. 482-491.

[31] Li, H., D. Arditi, and Z. Wang, Transaction-related issues and construction project performance. *Construction Management and Economics*, 2012. 30(2): p. 151-164.

[32] Terziovski, M. and J.-L. Guerrero, ISO 9000 quality system certification and its impact on product and process innovation performance. *International Journal of Production Economics*, 2014. 158: p. 197-207.

[33] Khalil, N., S.N. Kamaruzzaman, and M.R. Baharum, Ranking the indicators of building performance and the users' risk via Analytical Hierarchy Process (AHP): Case of Malaysia. *Ecological Indicators,* 2016. 71: p. 567-576.

[34] Mirza Jahanzaib, D. and K. Akhtar, Prioritizing Key Quality Indicators for Engineering Education using Analytical Hierarchical Process (AHP): Case Study, in *Proceedings of Pakistan's 4th National Conference on Quality Assurance in Education (NCQAE) Lahore*- September 17-18, 2005. 2005: Lahore.

[35] Mansor, M.R., et al., Hybrid natural and glass fibers reinforced polymer composites material selection using Analytical Hierarchy Process for automotive brake lever design. *Materials & Design*, 2013. 51: p. 484-492.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Effects of Economic and Population Factors on Health Expenditures: Special Case of Pakistan

Tanzeela Yaqoob, Rahat Bibi and *Junaid S. Siddiqui
Department of Statistics, University of Karachi

***Abstract*- This paper aims to explore the factors affecting the health care expenditures of Pakistan by employing the Multivariate techniques for the annual data series from 1960 to 2010 .The variables we considered are Gross Domestic Product (GDP) as an economic indicator, Population of age 65 and above (P≥65), Population of age 0 to 14 (P(0-14)), Life expectancy at Birth (LE), Crude Birth rate (CBR) and population growth rate (GR) as population indicator and Total Health Expenditure (THE) as an influencing factor. Multiple regression considering Total Health Expenditure (THE) is applied as dependent on the variables mentioned above. Since Variance Inflation Factor (VIF) for all independent variables are very high, the smallest VIF is around 11 goes up to 65. The results obtained by Principal components reveal the effect of population structure and the age factor effect oppositely on Total Health Expenditures (THE). Factor Analysis suggests that the behavior of Health Expenditures is common with Gross Domestic Product (GDP), population of age 65 and above (P≥65) and life expectancy at Birth (LE) in Pakistan while population of age 0 to 14 (P (0-14)), population growth rate (GR) and Crude Birth rate (CBR) are moving in opposite direction i.e. as GDP, P≥65 and LE of Pakistan increases P (0-14), GR and CBR will decreases.**

Keywords: Health care expenditure, principal Component Analysis, Factor Analysis.

I. Introduction

Since the seminal work of Joseph P. Newhouse [1] to explore the relationship between per capita health spending and per capita Gross Domestic Product (GDP) of 13 developed countries, the new horizon of research has been opened. Over the past few years the determinants of health expenditure have been an attractive topic for health economists. Inspiring by the numerous studies in this context this study attempts to study the effects of economic and population indicators on health expenditure for the emerging country Pakistan. As noted by khan et al. [2] the income elasticity with respect to health care expenditure is less than one thus in the developing countries health related aspects are seems to be necessity rather than luxury. The rising population and poor health facilitation provided by government may cause high mortality rate in Pakistan. Akram et al. [3] suggested that for sustainable economic growth the special attention must be given to health and education sectors of Pakistan.

In Pakistan health care is one of the most important issues of the country, but unfortunately the

*jsdr123@yahoo.com

percentage of GDP invested in this particular problem is alarming. The special attention should be given to financing system of Pakistan towards health sector. Out of the five methods of financing the health care system, Pakistan only utilizes general tax revenue and out of pocket payments. According to WHO [4] Pakistan ranks at high number among the highest out of pocket expenditure as percentage of private expenditure on health. The Total health expenditure as percentage of GDP in India, Bangladesh and Malaysia recorded by WHO are 4.1, 3.4 and 4.4 in 2007 respectively and Pakistan has the minimum of all i.e. 2.7 percent in the same year.

The beginning of 20th century brings the developing countries into the healthy world by innovating the health facilities, improved food and nutrition, better health infra structure and developed technology. These all can make possible by allocating the large share of budget in health sector by government as well as implementing the sound strategies in the country by policy makers to ensure the facilities of health. This is the reason that voluminous literature has been focused to link the health expenditure and economic growth for emerging economy countries.

Growing population may be the additional factor for middle income countries like Pakistan to allocate the health care facilities uniformly among the population. Crude Birth Rates of more than 30 per 1000 are considered high and rates of less than 18 per 1000 are considered Low. Since the beginning of the 90's, Pakistan has battled between the two numbers, from being too close to the high crude birth rate at the start of the 90's, with 26 births per 1000 in a population; it saw a gradual decline in the rate. By the end of the decade, Pakistan had a Crude birth rate of almost 22, and the decade's average amounted to 23 births per 1000 people which indicate a substantial increase in the population of Pakistan.

At the start of the last decade of the 21st century, Pakistan had only diverted, on average, a mere 5% of its GDP to health care, and this trend continued throughout the 90's. However, at the beginning of the 21st Century, the GDP percentage for health was increased but that increase was nothing when compared to the mortality rates that plagued the country. In 2000, Pakistan allocated 5.22 percent of its GDP to health which quite simply not enough. As the decade progressed we can see an upward trend with the percentage rising to 5.44, 5.63, and 5.77 in 2001, 2002, and 2003, respectively. The years that followed saw an already inadequate percentage decline even further as it reached a dismal 5.65 percent within a 4-year period. When the economies were being strained by the Market Collapse in 2008, the percentage was remarkably increased and by the 2009 it had reached a record-breaking 6.32% in 2009. But unfortunately, where it took several year for it to reach that point, it didn't take long for it to move back down and within a 5-year period the percentage of health-care had declined to a mere 6 percent by the end of 2014.

This study undertakes empirical analysis to understand the patterns that how health expenditure, population indicators and economic indicator is related. Numerous studies have been focused to identify the determinants of health expenditure by using different statistical and economical

models; this study aims to observe that how population and economic indicators effect total health expenditure in Pakistan by applying the multivariate techniques to avoid the multi-collinearity and heteroscedasticity problems in time series data. The remainder paper is designed as follows: The second section explains brief overview of literature related to the study, remaining sections describes the data sources and summary statistics, methodology, Empirical results and discussion followed by conclusion.

II. Review of Literature

To access the relationship between Health Expenditures and Economic Growth in Middle East & North Africa (MENA) Countries, Panel cointegration technique is utilized by Mehrara et al. [5]. They collected the annual data for 13 MENA countries from various sources spanning from 1995 to 2005. After examining the time series properties of Health Expenditure and Gross Domestic Product they employed traditional Panel cointegration technique and shows the long run co integrating equilibrium relationship holds between Health Expenditure and Gross Domestic Product. Similarly, they also conclude that Health care is necessity rather than a luxury good.

AK [6] explores the relationship between health expenditures, economic growth (GNP) and life expectancy at birth for Turkey. The long term causality relationship between the said variables have identified by Johansen cointegration technique while short term dynamics have been investigated by employing error correction model. The study concludes the long run existence of relationship however short run relationship between the variables is not significant.

SAARC countries from 1995-2012 have been analyzed to examine the Health care expenditure and economic growth by Khan et al. [7]. They include Per capita income, labor force, literacy rate, and elderly population of age 65 and above as a set of independent with Health care expenditure as dependent variable respectively. Similar to various studies they investigate short run as well as long run relationship by introducing panel Dynamic Ordinary Least Square (DOLS) and Seemingly Unrelated Regression (SUR) techniques. The study reveals long run relationship while the short run bi-directional causality exists between the variables.

Hitris [8] studied the 10 EC member states: Belgium, Denmark, France, Germany, Greece, Italy, Ireland, the Netherlands, Spain and the United Kingdom. The annual data from 1960-1991 has taken for selected variables. The log-linear model was utilized to find the relationship of per capita health expenditure both private and public with per capita GDP, Population $\geq$ 65 share of health care expenditure in total public spending, the rate of inflation and dummy variable to account for differences in countries health service systems. The study concludes that health expenditure depends in different way for each country as GDP and health care system are different.

Chaudhry et al. [9] investigates the factors related to health effects the economic growth of

Pakistan. The data on variables like gross fixed capital formation, employed labor force, exchange rate, inflation rate, total number of beds in hospitals, health expenditures, total number of dispensaries and total number of registered doctors have collected from 1975 to 2010 and employed autoregressive distributed lag model (ARDL) to accomplish the required objective. The long run model concludes that in Pakistan economic growth is improving due to the improvement in health facilities similarly inflation rate increases as economic growth increases while in short run inflation and growth is significant but negatively related.

Haider and Butt [10] utilize the data of annual time series from 1972 up to 2005. Their main objective is to enquire the nature of relationship between health care spending and Gross Domestic Product (GDP). After applying the time series technique of vector auto-regression (VAR) and autoregressive distributed lag model (ARDL), they conclude that elasticity of demand of health care is greater than unity so health care is necessity instead of luxury for the emerging country like Pakistan and long run relationship is evident between the economic growth and health care spending.

The positive relationship between health and economic growth and negative relationship between life expectancy and economic growth of SAARC countries is observed in the study of Zafar [11]. Panel EGLS method is applied on the yearly data of SAARC countries extending from 1986 to 2010. Health, GDP, fertility rate, life expectancy is incorporated in the study and theoretically makes linkage among them and showing that when health increases life expectancy increases and mortality rate decreases, when life expectancy increases mortality rate decreases and human capital increases, similarly increase in production yields high exports and ultimately economy grows. Thus study concludes that in order to get stable economy, SAARC countries put efforts towards better health facilitation.

Devlin and Hansen [12] finds health care expenditure granger cause Goss Domestic Product (GDP) and inverse is also true for 20 OECD countries from 1960 till 1987. Highlighting the importance of misspecification in the model authors suggest special care should be taken for the selection of suitable exogenous and endogenous variable in order to find the relationship between health and economic growth. The study overall concludes the positive relationship among the two said variables and more theoretical and empirical research has been suggested by the authors.

By using the per capita GDP as dependent variable and age dependency, openness, population per bed, secondary school enrolment, life expectancy, mortality rate and health expenditure as an independent variable, Akram et al. [3] explore that health expenditure and per capita GDP are not related while other independent variables influence the dependent variable. The study reveals that there is no conitegration between health care expenditure and per capita GDP in Pakistan. The cointegration technique followed by Error correction model is studied to find long and short run analysis respectively. The study concludes that health influences economic indicator only in long run.

Ozturk and Topcu [13] investigate the relationship between health care and economic growth among G8 countries i.e. developed nations in the world. Panel error correction model is applied to the secondary data ranging from 1995 to 2012 for each country. The study concludes that in short run the uni-direction causality exists from health expenditure to economic growth while the opposite direction of causality reveals in long run.

Bloom et al. [14] employs the production function approach to investigate the effect of health in labor productivity over the panel data of 104 countries for the time period from 1960 till 1990. Using the life expectancy as health indicator and applying the econometric approach of two stage least square they conclude that life expectancy and schooling are significant and positively influence on economic growth. Similarly, they also find that 4 percent of output is increasing due to the one-year enhancement in life expectancy.

Granger causality test is used to find the role of human capital in economic growth of Pakistan by Jangraiz Khan et al. [2]. For human capital they include research and development (R&D), education and health and for economic growth per capita GDP is incorporated in the analysis. The study reveals a unilateral causal relation between the variables while research and development has a causal relationship with economic growth. Furthermore, study concludes that life expectancy granger cause education.

III. Methodology

Principal component analysis (PCA) is one of the oldest techniques examined by Karl Pearson [15] while Hotelling [16] explains the mathematical procedure for computing it. It is the simplest technique to carry out as it does not require any particular assumption to fulfill before analyzing the data. The analysis required to take n variables

$x_1, x_2, x_3, \ldots\ldots\ldots\ldots x_n$

and search a composition of these variables to produce uncorrelated indices

$y_1, y_2, y_3, \ldots\ldots\ldots\ldots y_n$

These linear combinations are ordered in such a way that captures the highest amount of variation in it, captures the second highest mass of variation and so on. The objective of creating these linear combinations is to extract such components whose variances is large among all components, thus neglecting the low variance components we extract those one which effectively describe the variation in data set.

The procedure of principle component analysis consists of following steps; Starting with the set of n variables for p individuals, extract i^{th} principal components expressed as linear combinations of

original variables $x_1, x_2, x_3, \ldots\ldots\ldots\ldots x_n$

,

$$y_1 = b_{11}x_1 + b_{12}x_2 + b_{13}x_3 + \ldots\ldots\ldots + b_{1n}x_n$$

$$y_2 = b_{21}x_1 + b_{22}x_2 + b_{23}x_3 + \ldots\ldots\ldots + b_{2n}x_n$$

.
.
.

$$y_i = b_{i1}x_1 + b_{i2}x_2 + b_{i3}x_3 + \ldots\ldots\ldots + b_{in}x_n$$

These components varies as much as possible subject to the constraint that

$$b_{11}^2 + b_{12}^2 + b_{13}^2 + \cdots + b_{1n}^2 = 1$$

$$b_{21}^2 + b_{22}^2 + b_{23}^2 + \cdots + b_{2n}^2 = 1$$

.
.
.

$$b_{i1}^2 + b_{i2}^2 + b_{i3}^2 + \cdots + b_{in}^2 = 1$$

These constraints must hold as to maintain that also that is uncorrelated to while is uncorrelated to and . The method also includes calculation of the eigenvalues of the symmetric sample covariance matrix i.e.

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{bmatrix},$$

Where diagonal elements present the variances of original variables and off diagonal elements are covariance of variables. The eigenvalues are in fact variances of the principal components and they sorted as

$$\omega_1 \geq \omega_2 \geq \omega_3 \geq \cdots \geq \omega_i \geq 0$$

Now discard the components capturing low proportion of variation in the data i.e. extract those one holding 90% of the variation of the data set.

Spearrman [17] developed the idea of factor analysis by studying the correlations between the test scores of several types. There are three main stages required to perform factor analysis. Starting with the same aim of principal component analysis, n variables is taken for analysis and extracts p

provisional factor loadings and their uncorrelated linear combinations are constructed. In the second stage factor rotation is applied to make provisional factors in interpretable form, this rotation can be orthogonal or oblique. Orthogonal factor rotation that is mostly used is varimax rotation. Finding the factor scores leads the analysis to the final stage.

IV. Data Description And Summary Statistics

Annual data series spanning from 1960 till 2010 is examined. The data has been collected from various bulletins published by Economic Survey of Pakistan and State Bank of Pakistan. The variables Population of age 65 and above, (P≥65), Population of age 0 to 14 (P(0-4)), Life expectancy (LE), Crude Birth rate (CBR) and population growth rate (GR) are taken as the representative of population, Gross Domestic Product (GDP) as an economic indicator and Total Health Expenditure (THE) in Millions are included in the study.

The Visual display of population, economic indicators and Health expenditure as % of GDP and in million rupees has displayed in figure 4.1. The left panel (a) shows the annual GDP in Millions which indicates the gradual increase after the 1990's in Pakistan and it remains increasing at the end of sample period. Thus economic growth of Pakistan becomes stable after 2000.The right panel (b) of Figure 1 shows the Health Expenditure as % of GDP, indicating after 1985 it is declining and this pattern is observing till 2010. It can be seen from (c) that Crude Birth rate is in declining trend, since entering the 21st Century, this birth rate has been declining, a major reason for which is the increasing infant mortality rate rather than controlled births. With the dismal health care provided by the Government and a below-par medical service in the public sector, this crude birth rate had decreased even further and as the first decade of the 21st Century came to an end, Pakistan has come dangerously close to the Low-crude-birth-rate mark with the crude birth rate standing at an alarming 19 births per 1000 people while life expectancy of Pakistan is flourishing after entering into the 29th century. Young population (P (0-14)) is decreasing due to increment in mortality rates while old population (P≥65) is increasing as depicted in (e) and (f) respectively. Moreover, Population growth rate is declining as shown in (g), while (h) presents the Total Health Expenditure in million rupees indicating gradual increase as GDP increases after 1990.

The summary statistic of each variable is calculated and results are reported in Table1. It can be seen that GDP and THE showing leptokurtic behavior which can be also depicted in figure1 (a) and (h). Moreover, the mean and median of population structure i.e. P (0-14), P≥ 65, LEB, PGR and CBR are quite similar.

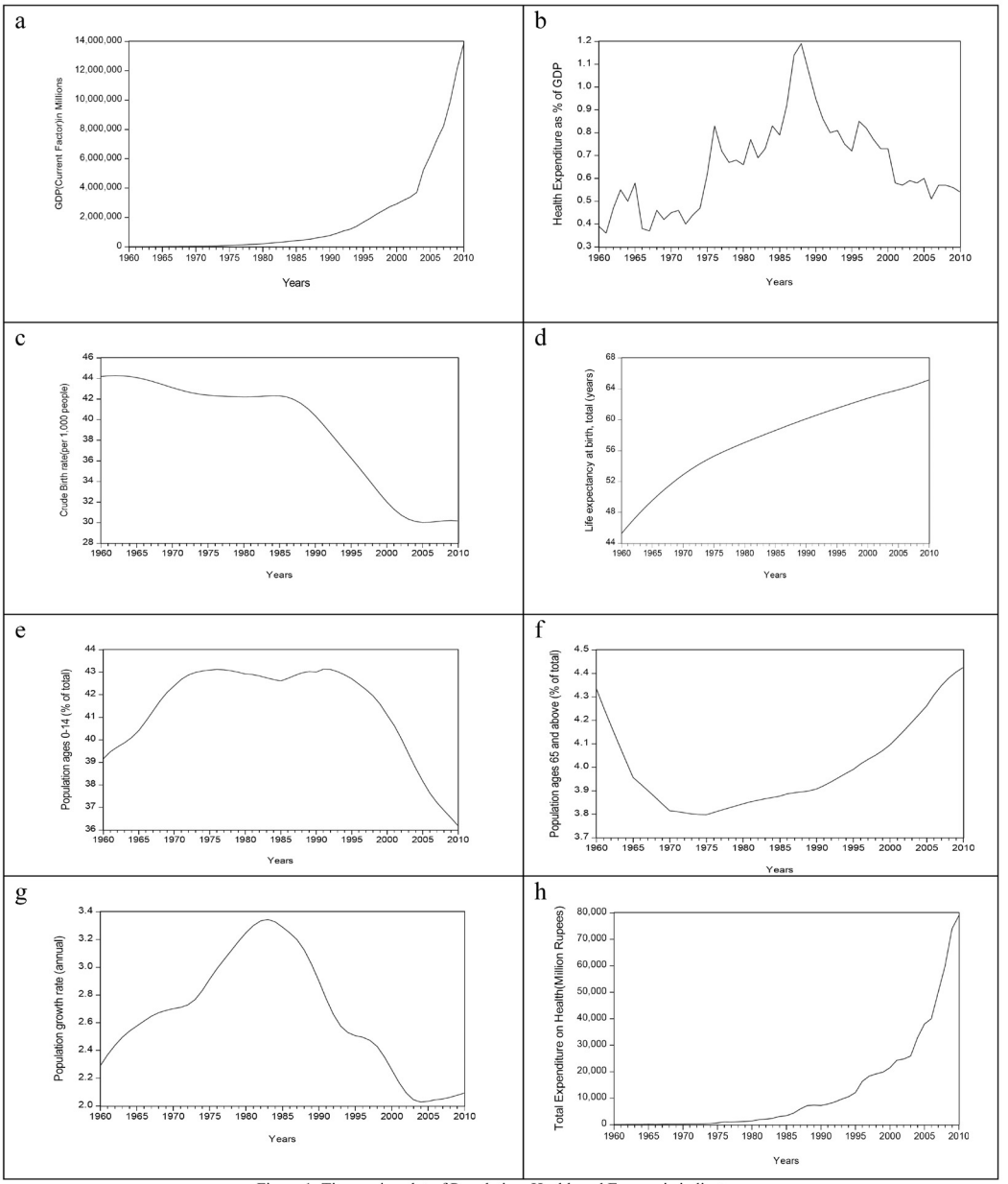


Figure 1: Time series plot of Population, Health and Economic indicators

TABLE 1
SUMMARY STATISTICS OF POPULATION AND ECONOMIC INDICATORS AND HEALTH EXPENDITURE

Statistic	GDP	THE	P (0-14)	P≥ 65	LEB	PGR	CBR
Mean	1916931.8	12663.64	41.39211	4.001226	57.66185	2.65543	38.96822
Standard Error	448341.18	2671.145	0.284789	0.02611	0.761007	0.058341	0.7365
Median	425064	3372.12	42.4761	3.92951	58.61807	2.646583	42.204
Standard Deviation	3201796.5	19075.79	2.033798	0.186464	5.434678	0.41664	5.259664
Sample Variance	1.025E+13	3.64E+08	4.136334	0.034769	29.53573	0.173589	27.66406
Kurtosis	5.0126998	4.005002	0.236251	-0.43154	-0.57262	-1.07688	-1.10773
Skewness	2.3017383	2.061201	-1.16388	0.874688	-0.60951	0.101011	-0.76088

To find the strength of relationship among the variables, correlations have obtained for the variables. Table 2 presents the results; absolute value of correlations between the variables varies from 0.27304 to 0.994874. The high negatively correlated pairs of variables are GDP with P (0-14), THE with P (0-14), P (0-14) with P≥ 65, THE with CBR, GDP with CBR, and CBR with LEB. Similarly, the high positively correlated variables are P≥ 65 with GDP, THE with GDP, P≥ 65 with THE, PGR with P (0-14).

TABLE 2
CORRELATIONS BETWEEN POPULATION AND ECONOMIC INDICATORS AND TOTAL HEALTH EXPENDITURE

Variables	GDP	THE	P (0-14)	P≥ 65	LEB	PGR	CBR
GDP	1						
THE	0.994874	1					
P (0-14)	-0.80579	-0.78319	1				
P≥ 65	0.79645	0.798556	-0.92965	1			
LEB	0.654691	0.700966	-0.1885	0.324059	1		
PGR	-0.65649	-0.66916	0.792999	-0.83112	-0.27304	1	
CBR	-0.82604	-0.86378	0.5967	-0.7052	-0.84556	0.723251	1

To analyze the more complete picture of relation between population and economic indicators with Total Health Expenditure (THE), scatter plots has drawn in Figure 2.

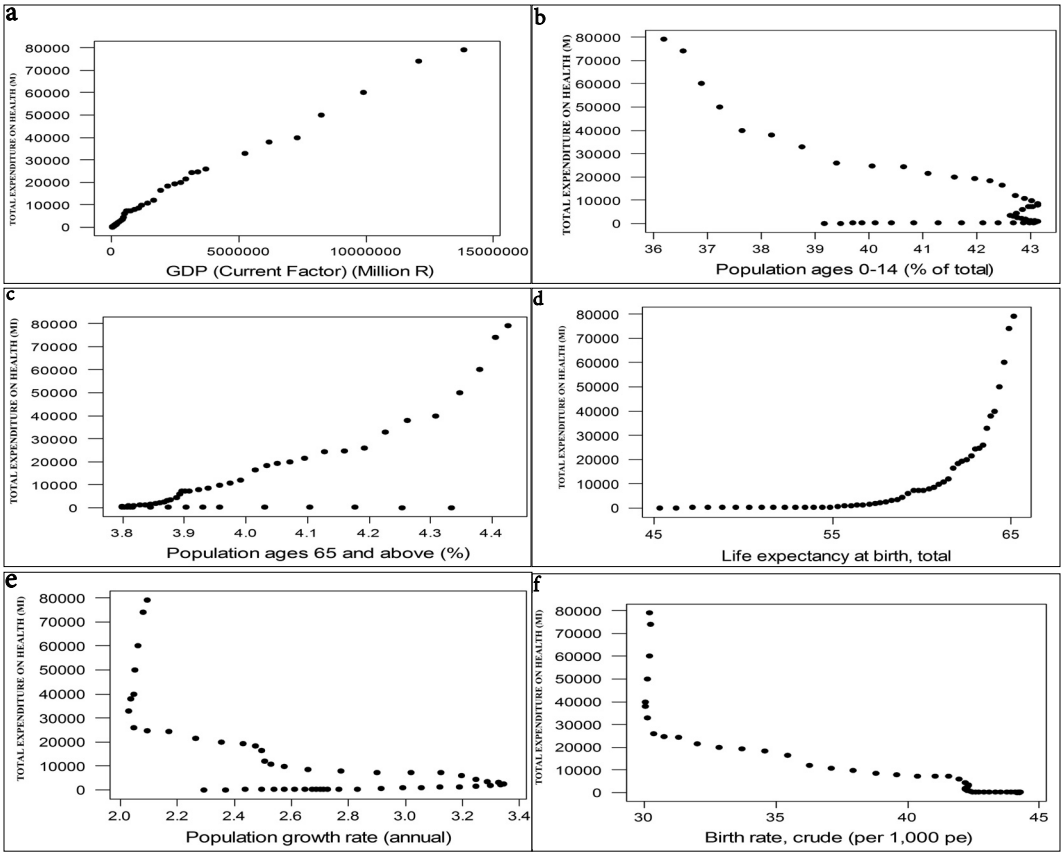


Figure 2: Time series plot of Population and Economic indicators Vs Health Expenditure

The figure 2 above clearly show the linear and positive relationship between old population (P≥65), and Gross Domestic Product (GDP) with Total Health Expenditure (THE). An exponential behavior is observing among Life Expectancy (LE) and Total Health Expenditure (THE).

V. Result And Empirical Analysis

To accomplish the required objective, the principal component analysis and Factor analysis on underlined variables is applied.

A. Principal Component Analysis

The first overview of the result has been obtained by excluding the total health expenditure and analyzes the behavior of variables. Table 3 presents the results in which the values less than 0.4 has been ignored. The Eigen values and their proportion also incorporate at the bottom of the table. We considered three PC's as they explain 97.7% of total variance of the observable data.

TABLE 3
PRINCIPAL COMPONENT ANALYSIS OF ECONOMIC AND POPULATION INDICATORS

Variables	**PCI**	**PCII**	**PCIII**
GDP	0.444	-	-0.523
P (0-14)	-0.415	-0.408	-
P≥ 65	0.437	-	-
LEB	-	-0.732	-
PGR	-0.408	-	-0.713
CBR	-0.433	-	-
Eigen value	4.379	1.1575	0.3248
Proportion	0.73	0.193	0.054
Cumulative	0.73	0.923	0.977

The first PC can be represented, approximately, by the following equation,

$$PCI = .4[(GDP+P\ (0\text{-}14)-(Pop\geq 65 + Growth\ Rate + CBR)]$$

The above expression represents contrast between GDP and young population with old population and growth of population. If young population and GDP (YPGDP) of the nation are moving together but the older population and causes of growth of population (OPGP) goes on the opposite direction of YPGDP. It explains 73% of variation.

$$PCII= 0.4[P\ (0\text{-}14) + 0.73(Life\ Expectancy)]$$

It is showing the factor of young population and life expectancy. This component contributes 19.3% of the total variation in the manifest data. We can code PCII as YPLE. We may conclude that 19.3% of variation in data structure is due to the fact of young population and life expectancy, therefore we are of the opinion that as young population increases the life expectancy is also increasing in Pakistan.

$$PCIII=0.523GDP+0.713Growth\ Rate$$

This component contributes only 5.4% of the total variation; we consider it because this is peculiar factors that GDP and growth rate moves along same direction. We code this PC as GDPGR. Similarly, we can see that the first three factors contain 97.7% of information.

These three principal components are considered as three independent variables and we will regress them on Total Health expenditure. This we are doing because when we regressed the observable variables GDP, P (0-14), P≥ 65, LE, PGR and CBR on Total Health expenditure the

value of R squared is very high i.e. 99.6% but variance inflation factor (VIFs) of the variables varies from 10.9 to 67.2. This makes the model redundant.

The following table 4 displays the result of regression analysis obtained by regressing the total health expenditure on the three extracted principal components named as YPGDP, YPLE and GDPGR respectively.

TABLE 4
ESTIMATION RESULTS OF REGRESSION OF TOTAL HEALTH EXPENDITURE AND EXTRACTED PCS

Predictor	Coef	StDev	T-Stat	P-value	VIF
Constant	12663.6	459.8	27.54	0	
YPGDP	8572.8	221.9	38.63	0	1
YPLE	-2936.1	431.7	-6.8	0	1
GDPGR	-8219.2	814.9	-10.09	0	1

We can see the results are quite robust now and the coefficients of PCs are highly significant and values of Variance Inflation Factor (VIF) are also in control mood. The following figure 3 portrays the plot between the obtained PCs.

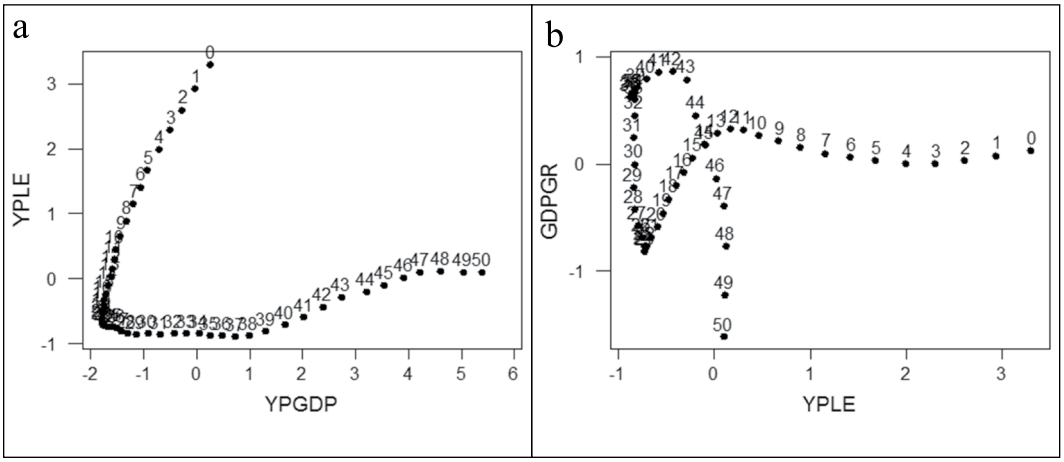


Figure 3: Plot of extracted Principal Components

The numbers showing on the plot represents the years starting from 0 as 1960 till 50 as 2010. The left panel (a) of figure 3 shows the Plot between First Principal component named as YPGDP and second component YPLE while the right panel (b) is showing the plot between YPLE, the second component and GDPGR the third component respectively.

The results extracted by adding the Total Health Expenditure in the analysis following output has been obtained.

TABLE 5
PRINCIPAL COMPONENT ANALYSIS OF ECONOMIC, HEALTH AND POPULATION INDICATORS

Variables	**PCI**	**PCII**	**PCIII**
GDP	0.414	-	-0.432
THE	0.418	-	-
P (0-14)	-0.373	-0.43	-
P≥ 65	0.391	-	-
LEB	-	-0.7	-
PGR	-	-	-0.682
CBR	-0.396	-	-0.387
Eigen value	5.2853	1.1807	0.3766
Proportion	0.755	0.169	0.054
Cumulative	0.755	0.924	0.978

By analyzing the results presented in Table 5 we found that the first three PC's explains 97.8% of information after adding the Health Expenditure, thus we figure out the first three PC's and exclude the coefficients less than 0.35. The first PC can be represented by following expression;

$$PCI=0.4[(GDP+THE+ Pop\geq 65)-[P\ (0\text{-}14) + CBR)]$$

The above expression is showing the contrast between old age population i.e. Pop≥65 (OP), Gross Domestic Product (GDP), Total Health Expenditure (THE) and Young population i.e. P (0-14) (YP), Crude Birth Rate (CBR). We named this PC ac OPGDPTHE. The second PC can be written as;

$$PCII=0.4\ [P\ (0\text{-}14)] + 0.7[(Life\ Expectancy)]$$

The above expression of PCII is similar to the PCII obtained previously showing that as young population i.e. P (0-14) and Life Expectancy(LE) both moves on the same direction, one increases the other also increases. We quote this PC as YPLE.

Following is the third PC which we got;

$$PCII=0.4[(GDP+CBR)] + 0.6\ Growth\ Rate$$

This explains the combination of Gross Domestic Product (GDP), Population growth rate (GR) and Crude Birth rate (CBR), thus we call this PC as GDPGRCBR. This component contributes 5.45% of the total variation in the observed variables.

Following Table 6 shows the regression results by regressing the total health expenditure on the three extracted principal components named as OPGDPTHE, YPLE and GDPGRCBR respectively.

TABLE 6
ESTIMATION RESULTS OF REGRESSION OF TOTAL HEALTH EXPENDITURE AND EXTRACTED PCS

Predictor	Coef	StDev	T	P	VIF
Constant	12663.6	306.6	41.31	0	
OPGDPTHE	7977.5	134.7	59.24	0	1
YPLE	-2456.9	284.9	-8.62	0	1
GDPGRCBR	-6496.7	504.5	-12.88	0	1

We can see from the estimated results that first pc (OPGDPTHE) effects positively on Total Health Expenditure (THE) while second (YPLE)and third effects (GDPGRCBR) negatively on Total Health Expenditure (THE).

B. Factor Analysis

The factor analysis has been analyzed first by taking all variables and the results have gathered in the Table 7.We consider the first three factors as they captures 97.8% of the total information contains in the observed data.

TABLE 7
ROTATED FACTOR LOADINGS OF ALL VARIABLES

Variables	FI	FII	FIII
GDP	0.774	-0.57	-
THE	0.728	-0.619	-
P (0-14)	-0.842	-	-0.516
P≥ 65	0.725	-	0.607
LEB	-	-0.988	-
PGR	-	-	-0.883
CBR	-0.396	0.795	-0.497
Eigen value	2.6444	2.4017	1.7965
% of Var	0.378	0.343	0.257

Results present the loadings of three extracted factors. The first factor model explains very well most of the original variables describing the overall economic and population indicators along with the health expenditure. The second factor is negatively related to Gross Domestic Product (GDP), Total Health Expenditure (THE) and Life Expectancy (LE) and positively related to Crude Birth Rate (CBR). The third factor is related mainly to the population indicators i.e. Young population (P (0-14)), Old population (P≥ 65), Population Growth Rate (GR) and Crude Birth Rate (CBR). Analyzing the scatter plot of first two factors, we find that young population, population growth rate and crude birth rate are moving together while the remaining variables somehow are on same path.

The results of factor analysis by excluding Total Health Expenditure (THE) are obtained in the following Table 8.

TABLE 8
ROTATED FACTOR LOADINGS

Variables	FI	FII	FIII
GDP	-0.765	0.596	-
P (0-14)	0.913	-	-
P≥ 65	-0.813	-	-0.479
LEB	-	0.993	-
PGR	0.508	-	0.830
CBR	-	-0.798	0.463
Eigen Values	2.4867	2.0716	1.3031
% of Var	0.414	0.345	0.217

After the preliminary examination three factors has extracted. Using the threshold of 0.4, we neglect the factor loadings lower than the 0.4. The first factor corresponds to population indicators i.e. young population (P (0-14)), old population (P≥ 65) and population growth rate (PGR). The second factor seems to be positively related to Gross Domestic Product (GDP) and Life expectancy (LE) and negatively related to Crude Birth Rate (CBR). The third factor combines old population (P≥ 65), population growth rate (PGR) and crude birth rate (CBR).

Figure 4 presents the scatter plot of three extracted factors. The left panel (a) of figure 4 drawn between Factor I and Factor II showing that Gross Domestic Product (GDP) and old population (P≥ 65), Crude Birth Rate (CBR) and Population growth rate (GR) is moving together respectively. While the right panel (b) of figure drawn between Factor II and Factor III displaying that Population growth rate (GR) and Crude Birth Rate (CBR), old population (P≥ 65), Life expectancy (LE) and Gross Domestic Product (GDP)lies in the same line.

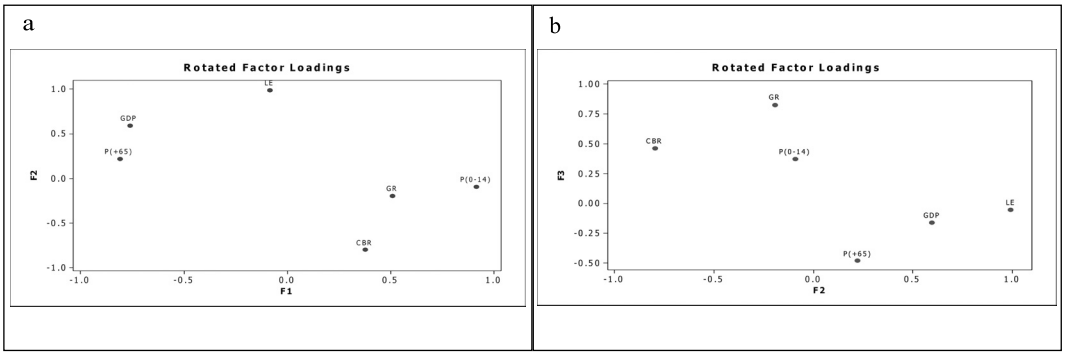


Figure 4: Rotated Factor Loadings plot

VI. Conclusion

The strength of development for any country hides behind its economic growth and the economic growth as presented by Chaudhry et al. [9] is influenced by the health sector. By considering the population indicators along with Gross Domestic Product (GDP) and Total Health Expenditure (THE), annual time series of the selected variables has taken to employ Principle Component Analysis (PCA) and Factor Analysis (FA). In developing countries where people engaged more in labor work, they require healthier mind and body to maintain their hourly wages as healthy worker perform in better way. Bloom et al. [14] emphasizes to improve the health status of country in order to improve the economic condition as healthier person can do more work as compared to an ill person.

Results from PCA reveal that the effect of population structure, Gross Domestic Product (GDP) and Total health expenditure (THE) comprises together in the same direction indicating that the special attention should be given towards health sector as to improve economic stability and also to maintain the population structure of the country. Factor Analysis suggests that the behavior of Health Expenditures is common with Gross Domestic Product (GDP), population of age 65 and above and life expectancy in Pakistan while population of age 0 to 14, population growth rate and Crude Birth rate are moving in opposite direction i.e. as GDP, P≥65 and LE of Pakistan increases P (0-14), GR and CBR will decreases.

This study can further be extended by adding the per capita income of Pakistan in order to see the broader picture while discrimination of health facilities between urban and rural areas must be incorporated to visualize the real status of health in Pakistan. Economist and policy makers should take immediate steps to enhance the health sector in Pakistan not only by establishing hospitals or increasing the number of doctors but also emphasize on creating awareness regarding the health problems of an individual.

References

[1] Newhouse, J.P. (1977). Medical-Care Expenditure: A Cross-National Survey, Journal of Human Resources, vol.12, pp, 115-125.

[2] Khan, J.,Khattak,N.U.R.,Khan,A. (2015). Human Capital-Economic growth Nexus: A causality analysis for Pakistan, City University Research Journal, vol.5, pp, 279-290.

[3] Akram,N., Padda,I.U.H., Khan,M. (2008). The long term impact of health on Economic growth in Pakistan, The Pakistan Development Review, vol.47, pp, 487-500.

[4] World Health Organization.(2010). World health statistics (pp. 130–139).Geneva, Switzerland.

[5] Mehrara,M., Fazaeli,A. A., Fazaeli,A.A., Fazaeli ,A.A. (2012). The Relationship between Health Expenditures and Economic Growth in Middle East & North Africa (MENA) Countries, International Journal of Business Management and Economic Research, vol.3, pp ,425-428.

[6] AK, R. (2012). The Relationship between Health Expenditures and Economic Growth: Turkish Case, International Journal of Business Management and Economic Research, vol.3, pp, 404-409.

[7] Khan,H.N., Khan,M.A., Razli,R.B., Sahfie,A.B., Shehzada,G., Krebs,K.L., Sarvghad,N. (2016). Health Care Expenditure and Economic Growth in SAARC Countries (1995–2012): A Panel Causality Analysis, Applied Research Quality Life, vol.11, pp, 639-661.

[8] Hitiris,T. (1997). Health care expenditure and integration in the countries of the European Union, Applied Economics, vol.29, pp, 1-6.

[9] Chaudhry,M.O., Faridi,M.Z., Farooq,F.,Arif,R. (2013). Contribution of Health Outcomes to Economic Growth in Pakistan, Pakistan Journal of Social Sciences, vol. 33, pp, 281-295.

[10] Haider,A., and Butt,M.S. (2007). The Direction of Causality between Health Spending and GDP: The Case of Pakistan, Pakistan Economic and Social Review, vol.45, pp, 125-140.

[11] Zafar.,N. (2013). Does Improvement in Health Leads toward Economic Growth:A Case of SAARC Countries, Journal of global and science, vol.1, pp,25-29.

[12] Devlin,N. and Hansen,P. (2001). Health care spending and economic output: Granger causality, Applied Economics Letter, vol. 8, pp, 561-564.

[13] Ozturk,S., and Topcu,E. (2014). Health Expendıtures and Economıc Growth: Evıdence from G8 Countrıes, International Journal of Economics and Empirical Research, vol.2, pp, 262-267.

[14] Bloom,D.E., Canning,D., Sevilla,J. (2004). The Effect of Health on Economic Growth: A Production Function Approach, World Development, vol.32, pp, 1-13.

[15] Pearson, K. (1901). On lines and planes of closest fit to a system of points in space, Philosophical Magazine, vol.2, pp, 557-572.

[16] Hotelling, H. (1933). Analysis of a complex of statistical variables into principle components, Journal of Educational Psychology, vol.24, pp, 417-441 and 498-520.

[17] Spearman, C. (1904). 'General intelligence', objectively determined and measured, American journal of Psychology, vol.15, pp,201-293.

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 6, No 2, December 2016

Bankruptcy Prediction for Financial Sector of Pakistan: Evaluation of Logit and Discriminant Analysis Approaches

*Usama Ehsan Khan

***Abstract*- Bankruptcy prediction is one of the core area in finance that is quite rich in empirical and theoretical work. This study compares two models for measuring the financial position of financial firms listed in Karachi Stock Exchange. The study gives a comprehensive review of two models, namely Altman's [1] Z-score and an O-Score derived from Ohlson [14]. The purpose of this paper is two folded. First to identify unique characteristics of business failure and to compare effective variables responsible for this response. Secondly to compare two popular accounting-based measures. summarize publicly-available information about bankruptcy. The sample period for this study is from 2009 to 2015. From the KSE listed financial firms, a total of 40 firms were selected and accounting ratios were extracted from balance sheet analysis reports published by State bank of Pakistan. The empirical results concluded that the logit model has a high rate of classification as compared to multiple discriminant analysis. The model has obtained overall 85.5% accuracy and identified three significant accounting ratios that are: retained earnings to total asset, earnings before income and taxes to the total asset, and current liabilities to total asset. The finding of this study would benefit stakeholders that are affected by bankruptcies. So in order to take an advantage, it is important to understand the phenomenon that causes bankruptcies.**

Keywords: Multiple Discriminant Analysis, Logit, and Z-Score.

I. Introduction

Financial distress is the situation when a company becomes unable to meet its financial obligations. There have been numerous studies regarding bankruptcies since the pioneering studies of Beaver [7] and Altman [1]. Multiple models and techniques have been proposed widely to examine bankruptcy in order to help investors, managers, and financial analysts. A prediction model can also be used to aid these decisions and that can act as a warning system. Events, like Global Financial crisis of 2007 challenged the predictability of the existing models. Predicting financial distress of companies from financial ratios have been explored in various ways and model may still be very useful. Extensive literature is available on predicting bankruptcy based on historical accounting ratios, security market information or macroeconomic indicators. Recent studies argued for the combined approach. Tinoco and Wilson [16] tested all three set of variables and concluded that the utility of combined approach is higher. These models have been extensively applied to the companies through statistical procedures such as MDA, logit, hazard and more recently artificial neural network (ANN) models.

*usama.ehsan@gmail.com

The present study contributes to the existing literature. It identifies the merits of logit model over discriminant analysis (DA). This model also helps in determining potential factors that are affecting significantly the financial status of the companies in the financial sector. Moreover, there has been no investigation of bankruptcy for the financial sector of the Pakistan, although the economy heavily relies on this sector. None of the world wide studies have particularly focused on Pakistan. There are only a few studies that have explored the financial vulnerability of the firms in the context of Pakistan but they are mainly concerned with non-financial sectors. The previous research of Abbas and Rashid [5] employed multiple discriminant analysis (MDA) for the non-financial sector of Pakistan and achieved 76.9% prediction accuracy. Mehta [13] conducted a study for assessing the financial position of textile sector. His study use 37 bankrupt and 53 non-bankrupt companies and modeled bankruptcy using logit discriminant model. Their model achieved 92% accuracy. Waqas, Hussain and Anees [17] used probit model and attained a maximum of 73.8% accuracy in classification. His study was based on 412 companies of the manufacturing sector of Pakistan.

Bankruptcy study is of great significance for practitioners, financial analyst and academicians. Pakistan is a developing country with the status of emerging markets. A large number of bankruptcy incidents have occurred in the last two decades, so there is a growing need of developing prediction models with a certain level of accuracy.

The structure of the rest of the paper is as follows. Section: 2 reviews the existing literature, section: 3 is a methodology and data, section: 4 show the results and section: 5 is the conclusion of the study.

II. Literature Review

Bankruptcy prediction and evaluation of financial distress are the core areas of finance. The literature is quite rich with bankruptcy prediction studies. Since the 1960s, researchers have been contributed much in order to examine the bankruptcy prediction all over the world.

The study of Beaver [7] is considered as the pioneering work in the area of bankruptcy prediction, which identified 30 ratios and by using univariate logistic model, the study found that best predictor was working capital to debt ratio and net income/total assets which correctly classified 90% of the cases. Altman [1] used a similar model for bankruptcy prediction by using multiple discriminant analysis (MDA) and developed Z - score model which identified five accounting ratios responsible for such discrimination. If the company's Z-score is above 2.675, it indicates a fine financial status; if it is less than 2.675 then it indicates that the company has financial vulnerabilities. Altman's [1] model remained a benchmark model for other studies. He revisited Z-score and ZETA model to improve the predictability by including non-manufacturing firms as well. This study was based on a sample of 66 corporations and found that model has a 95% rate of classification.

Ohlson [14] used a probabilistic model by using nine financial ratios and argued that probabilistic model has advantages over MDA as it does not require variables to be normally distributed. This is quite common that financial variables deviate from normal distribution. Recent study of Ahmadi, Soleimani, Vaghfi, and Salimi [6] also verified the superiority of logistic model over MDA in terms of bankruptcy prediction. Recently, Keener [11] showed the predictive ability of logit model by using a sample of 1203 retail firms and five financial ratios. His study concluded the extremely accurate prediction using logit. Hardle and Prastayo [10] showed the higher prediction accuracy of regularized logit model for South East Asian economies by using 24 accounting ratios. Their study employed Lasso and elastic-net penalty functions for the approach of regularization.

Several other studies also used different methods with different set of variables for assessing financial sustainability. For example, Cihak and Hesse [8], whose analysis examined the relative performance of Islamic banks and their financial status. Their study used 77 Islamic banks and 397 commercial banks from 20 different countries and by using z-score, the study concluded that small Islamic banks are financially healthy than small commercial banks, but the case is reversed for large banks. However, small Islamic banks tend to financially more stable than that of larger Islamic banks. Lennox [12] used a sample of 949 firms and examined the bankruptcy using logit, probit, and DA. The study revealed that well-specified logit and probit models can identify failing companies more accurately.

III. Data and Methodology

This section is devoted to the sample and the methodology employed in the study. It firstly tells about the sample and variable selection which is followed by the data analysis.

A. Sample and variable selection

The population consists of all those companies that are listed in KSE and the sample covers 20 financially sound companies and 20 financially distressed companies. Pindado, Rodrigues and Torre [15] stated two conditions that are needed to be met in order to predict the financial distress condition, (i) when earnings before interest, taxes, depreciation and amortization (EBITDA) are below their financial expense for two consecutive years, and (ii) firm suffers from negative growth in market value for two consecutive years. This study used following criteria for the financially distressed firm:

1) All financial companies that are delisted by the Karachi Stock Exchange (KSE) due to liquidation/ winding up under court order.

2) The company must have at least five years of financial information.

3) The company whose earnings before income and taxes, depreciation and amortization is

less than its expense for two consecutive years.

The total number of companies confirming the aforementioned criteria was 20, then 20 financially sound firms were randomly selected. So, there are total 40 firms with five years of observation available over the period of 2009-2015.

The data has been extracted from "Financial statement analysis of financial sector" published by the State bank of Pakistan. This study included following accounting ratios:

1) Working capital/Total assets (*WCTA)*. The working capital to total assets is the liquidity ratio found frequently in corporate problems. It is the measure of net liquid asset to the total capitalization. It is one of the five accounting ratios that were the best predictor classified by Altman [1].

2) Retained earnings/Total assets *(RETA)*. Retained earnings are the account that shows the total amount of reinvested earnings or loses over the entire life of the firm. It is also referred as earned surplus.

3) Earnings before income and taxes/Total assets *(EBITTA)*. It is the measure of the firm's productivity, refraining from any tax or leverage effect. Firm's survival is ultimately depending upon its earning capacity.

4) Book value/Total assets *(BVTA)*. Book value is also the net asset value. Book value to total asset is the fraction of the net asset from the total assets.

5) Debt to equity ratio *(DE)*. It is the ratio that compares the capital invested by owners to the capital provided by lenders. Since high debt to equity ratio can put the business at risk, so it is the appropriate ratio for the bankruptcy study.

6) Total income/Debt *(TID)*. TID proxy the leverage ratio. It proxies the long run leverage effect.

7) Current asset/current liabilities *(CACL)*. It is the ratio that ascertain whether a company's short term assets are available to pay-off its short-term liabilities.

8) Net income/Total asset*s (NITA)*. It is the most common ratio to measure profitability of the firm.

9) Current liabilities/Total assets *(CLTA)*. Current liabilities are short-term liabilities. CLTA is the measure of short-term leverage effect to the firm.

10) Cash/Current liabilities *(CCL)*. Cash available measure the liquidity of the firm and it is the measure component of current assets. CCL ratio is measuring liquidity of the firm.

The mean and standard deviation of ten financial ratios of the company are reported in the Table-1. It is evident that the high indebtedness, lower profitability and poor liquidity are the key

characteristics of the bankrupt firms which support our analysis.

TABLE 1
PROFILE ANALYSIS

Variable	Non-bankrupt firms		Bankrupt firms	
	Mean	Std. Deviation	Mean	Std. Deviation
WCTA	-.830	3.258	1.563	11.544
RETA	.170	.339	-.434	1.045
EBITTA	.052	.094	-.0388	.137
BVTA	6.051	24.523	11.145	29.428
DE	6.167	9.075	.560	27.242
TID	.622	2.301	-.838	7.427
CACL	8.888	33.568	42.480	146.155
NITA	299.695	1794.331	-.533	5.250
CLTA	20.278	81.003	4.849	23.826
CCL	7.194	33.772	14.257	78.996

B. Data Analysis Models

Over the last many decades bankruptcy prediction remained a hot topic. Extensive work has been done in this area and there are several models proposed by various studies of bankruptcy. Abbas and Rashid [5] obtained 76.9% accuracy by using MDA while Mehta [13] found 92% accuracy of logit model when applied to textile sector of Pakistan. This study models financial distress by using both MDA and logit model for the financial sector of Pakistan aimed to identify the model with higher rate of classification.

C. Multiple Discriminant Analysis (MDA)

Altman [1] derived a Z-score model for the bankruptcy prediction by combination of five ratios with the extension of Beaver's [7] approach. In 1977, Altman, Haldeman, and Narayan [2] obtained Z-score model for non-manufacturing firms that contains four accounting ratios. The model that is developed through MDA takes the following form:

$$Z= \beta_1X_1+ \beta_2X_2 +\beta_3X_3+ \beta_4X_4+\ldots.+ \beta nXn. \qquad (1)$$

Where Z is the overall Z-score and βi are the discriminant coefficients and X1, X2,…Xn are the independent variables. MDA technique can draw the best combination of the key ratios from the set of accounting ratios. We choose the best predictor by looking F and Wilk's lamda statistics to check

statistical significance of each individual ratio. Further, chi-square statistics are also used for- the overall significance of various discrimination functions.

D. Binary logit model

The profile analysis in Table-1 shows that the bankrupt firms have several unusual characteristics. In order to model the financial distress of the company through accounting ratios, an alternative approach is probabilistic model (i.e. logit). Following Ohlson [14], we now estimate the probability of default of the firm.

We assume that marginal distribution of the failure follows the logistic distribution of the following form:

$$\ln\left(\frac{pi}{1\text{-}pi}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \ldots + \beta n Xn \qquad (2)$$

IV. Empirical Results

This section provides the empirical estimates of multiple discriminant analysis (MDA) and logit model. Furthermore, receiver operating characteristic (ROC) curve technique is used to evaluate the predictive performance of the models.

A. Results of the discriminant analysis

A sample of 40 companies with five years of data resulted 200 firm-year observation. We initially used ten financial ratios and through step-wise multiple discriminant analysis (MDA) we identified six ratios that are depicted in Table-2. Ratios are selected according to their relative significance and maximum significance to enter the variable is 10%.

TABLE 2
Variables Entered / Removed

Step	Entered	Wilks' Lambda	F-Statistics	Sig.
1	*RETA*	.867	30.329	.000
2	*EBITTA*	.793	25.645	.000
3	*CACL*	.769	19.641	.000
4	*CLTA*	.752	16.113	.000
5	*WCTA*	.739	13.709	.000
6	*DE*	.727	12.050	.000

At each step, the variable that minimizes overall Wilk's Lamda is entered.
a Maximum number of step is 20
b Maximum significance of F to enter is 0.10.
c Minimum significance of F to remove is 0.20.

The discriminant procedure concluded significant variables and excluded the remaining variables as shown in Table-2. Consequently, from the set of prior ten accounting ratios, six accounting ratios are selected. Retained earnings to total asset (RETA), EBIT to total asset (EBITTA), current asset to current liabilities (CACL), current liabilities to total asset (CLTA), working capital to total asset (WCTA) and debt to equity (DE) were found significant at 1% level of significance.

The final Z score/discriminant score derived from Table-2, takes the following form:

$$Z=0.077 + 0.044WCTA + 1.110\ RETA+ 4.465\ EBITTA + 0.012\ DE – 0.003\ CACL + 0.005\ CCL$$

Where, Z is the Discriminant score. The midpoint or the cut-off value of the non-bankrupt and bankrupt centroid is zero. It can be concluded that the movement of the firm's Z-score above zero is approaching toward non-bankrupt group and the firm's Z-score below zero is approaching toward the bankruptcy. Finally, the firm having Z-score -0.609 classified as "bankrupt" and the firm having Z-score -0.609 is classified as "non-bankrupt", as depicted in Table-3.

TABLE 3
GROUP CENTROID

Group	Z-score
Non-bankrupt	.609
Bankrupt	-.609

The classification results presented in Table-4 compares the actual with the predicted outcome. The model has overall 81.5% accuracy of classification. It is evident that six significant financial ratios have high classification power which suggests that model has potential to predict corporate failure in Pakistan.

TABLE 4
CLASSIFICATION RESULTS: DISCRIMINANT ANALYSIS

		Predicted Group Membership		Total
	Z-score	Non-bankrupt	Bankrupt	
Original Count	Non-bankrupt	94	6	100
	Bankrupt	31	69	100
Percentage	Non-bankrupt	94.0	6.0	100.0
	Bankrupt	31.0	69.0	100.0

B. Results of the logit model

Logit model is very popular in prediction since it has a probabilistic interpretation embedded in it It can be seen from Table-5 that the signs of the estimates are almost same as expected. Coefficients in the table depicted are the odd ratios and the p-value is the probability value of the corresponding z-statistics. Three accounting ratios among the set of ten ratios are found significant i.e. Retained

earnings to the total asset (RETA), earnings before income and taxes to the total asset (EBITTA) and current liabilities to the total asset (CLTA).

Signs for RETA and EBITTA are negative as expected because firms with higher retained earnings to total asset ratio are less likely to default. Furthermore, EBITTA is the measure of the productivity of the firms so it was expected to have the negative relationship between probability of default and EBITTA which is supported from the empirical evidence of the study.

TABLE 5
LOGIT RESULTS

Variables	Coefficients	p-value
WCTA	0.033	0.726
RETA	-4.780*	0.000
EBITTA	-6.600**	0.022
BVTA	0.00008	0.993
DE	-0.018	0.263
TID	0.018	0.828
CACL	0.0093	0.277
NITA	-0.0013	0.819
CLTA	-0.0188*	0.009
CCL	-0.00929	0.381
_cons	0.1779	0.424

* Significant at 1%
** Significant at 5%

Classification results in Table-6 reported that out of 100 non-bankrupt firms-year observations, 91 were considered as non-bankrupt and from 100 bankrupt firms-year observation model successfully classified 80 firms as bankrupt so the model accuracy in this regard is 80%. Overall model acquired 85.5% prediction accuracy.

TABLE 6
CLASSIFICATION RESULTS: LOGIT MODEL

		Predicted Group Membership		Total
	Z-score	Non-bankrupt	Bankrupt	
Original Count	Non-bankrupt	91	9	100
	Bankrupt	20	80	100
Percentage	Non-bankrupt	91.0	9.0	100.0
	Bankrupt	20.0	80.0	100.0

Literature evident that the logit model has a high rate of classification of bankrupt and non-bankrupt companies. However in context to Pakistan, it was not revealed that whether it has a high rate of prediction in financial sector. This study shows that the logit model has potential to be used for bankruptcy in the financial sector.

C. Receiver Operating Characteristic (ROC) curve analysis

Receiver operating characteristics (ROC) curve was developed by engineers in world war-II in order to analyze the position of the enemy. Later it gained popularity in different fields and financial literature also made intensive use of ROC curve to compare the predictability of the models. Altman, Sabato, and Wilson [4] argued that the ROC curve plots the false positive rate and true positive rate to discriminate between two groups. Cleves and Rock [9] used ROC curve technique to compare the predictive accuracy of the probit and logit probabilistic models.

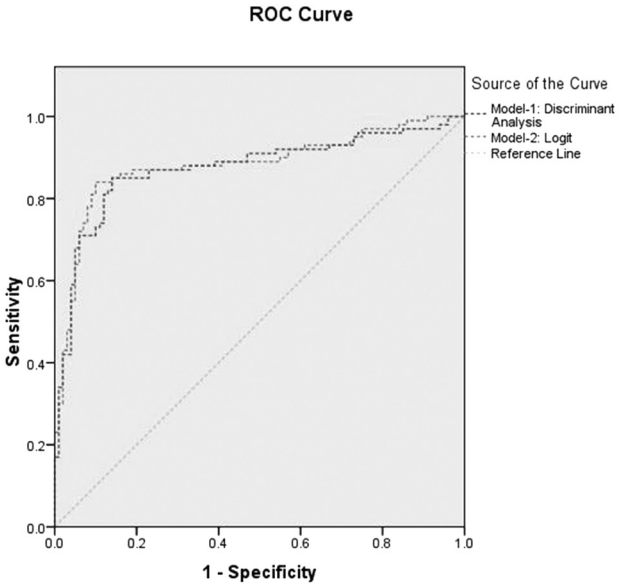


Figure-1: ROC curve analysis

The area under the ROC curve (AUC) corresponding to the model-1 is 87.4% and AUC corresponding to model-2 is 88.2%. Overall, AUC for the two models showed that model-2 corresponds marginally higher rate of classification.

V. Conclusion

Consistent with the previous studies, logit model has outperformed as compared to the discriminant analysis approaches. Whereas previous studies could not reveal the difference for

the financial sector of Pakistan. The financial sector is the backbone of the Pakistan's economy. Bankruptcy prediction work as a warning bell therefore should be accurate giving timely predictions. The discriminant analysis has produced a model with six accounting ratios namely retained earnings to total asset (RETA), EBIT to total asset (EBITTA), current asset to current liabilities (CACL), current liabilities to total asset (CLTA), working capital to total asset (WCTA) and debt to equity (DE). The model has attained an accuracy of 81.5%. The estimates provide that firms having a Z-score below zero are more toward bankruptcy, however, Z-score above zero indicates financially healthy firms.

In addition to this, logit model produced a parsimonious model with three significant financial ratios and has attained overall 85.5% accuracy. These three ratios are retained earnings to the total asset (RETA), earnings before income and taxes to the total asset (EBITTA) and current liabilities to the total asset (CLTA). It is evident that these three ratios are also among six ratios provided by DA. Furthermore, RETA and EBITTA were also identified by benchmark Z-score model of Altman [3].

In aggregate, this study provides comparison of the classification accuracy and predictive power of two well-known models in the area of bankruptcy. It contributes to the bankruptcy prediction literature using a broader definition of bankruptcy. The advantage is that the definition used in this study does not only correspond to the bankruptcy but includes those who are financially distressed. Second, this study revealed the bankruptcy situation in the financial sector, which has filled the gap in the existing literature. Third, it explored three significant variables that have a high predictive accuracy if modeled with logit. Furthermore, the study confirmed that logit model is the most parsimonious model as it possesses relatively high prediction accuracy with rather a small number of variables in the financial sector also. This study suggests that stakeholders should keep these significant ratios in order to assess the financial health of the firm.

References

[1] E. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," J. finance, 23(4), 589-609, 1968.

[2] E. Altman, G. Haldeman, and P. Narayanan, "ZETATM analysis A new model to identify bankruptcy risk of corporations," J. Banking and Finance, 1(1), 29-54, 1977.

[3] E. Altman, "Predicting financial distress of companies: revisiting the Z-score and ZETA models," Stern School of Business, New York University, 9-12, 2000.

[4] E. Altman, G. Sabato, and N. Wilson, "The value of non-financial information in small and medium-sized enterprise risk management," The J. Credit Risk, 6(2), 95, 2010.

[5] Q. Abbas, and R. Ahmad, "Modeling Bankruptcy Prediction for Non-Financial Firms: The Case of Pakistan," Available at SSRN 1917458, 2011.

[6] S. Ahmadi, B. Soleimani, H. Vaghfi, and M. Salimi, "Corporate bankruptcy prediction using a logit model: Evidence from listed companies of Iran" World Applied Sciences Journal, 17(9), 1143-1148, 2012.

[7] H. Beaver, "Financial ratios as predictors of failure," J. accounting research, 1966, pp.71-111.

[8] M. Čihák, and H. Hesse, "Islamic banks and financial stability: An empirical analysis," IMF Working Papers, 1-29,

2008.

[9] A. Cleves, and L. Rock, "From the help desk: Comparing areas under receiver operating characteristic curves from two or more probit or logit models," The Stata Journal, 2(3), 2002, pp. 301-313.

[10] K. Härdle, and D. Prastyo, "Default Risk Calculation based on Predictor Selection for the Southeast Asian Industry,"2013

[11] M. Keener, "Predicting The financial failure of retail companies in the united states," J. Business and Economics Research (Online), 11(8), 2013, pp. 373.

[12] C. Lennox, "Identifying failing companies: a re-evaluation of the logit, probit and DA approaches. J. Economics and Business, 51(4),1999, pp. 347-364.

[13] A. Mehta,"Rescuing Business; Analysis of Bankruptcy in textile Sector of Pakistan Using Logit Model From Year 2005-2010," Strategic Management Quarterly, June. 2014, pp. 81-99.

[14] A. Ohlson, "Financial ratios and the probabilistic prediction of bankruptcy," J. accounting research, 1980, pp. 109-131.

[15] J. Pindado, L. Rodrigues, and C.Torre, "Estimating financial distress likelihood," J. Business Research, 61(9), 2008, pp. 995-1003.

[16] H. Tinoco, and N. Wilson, "Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables," International Review of Financial Analysis, 30, 2013, pp. 394-419.

[17] H. Waqas, N. Hussain, and U. Anees, " Zmijewski financial distress prediction model and its predictability: A case of Karachi stock exchange. J. Basic and Applied Scientific Research, 4(4), 2014, pp. 155-163.