SECURITY AND TRUST IN ELECTRONIC DATA INTERCHANGE: A STUDY ON INDIAN AUTOMOBILE INDUSTRIES

K. K. Patra[1]

ABSTRACT

This paper intends to study the security aspects of electronic data interchange (EDI) services in Indian automobile industries. The paper has gone into great depth to address an EDI specific security requirement to the automobile sectors in India. EDI is the combination of techniques and technology, which facilitates paperless trading. It has resulted in the boosting of profit and productivity for business. While the paper introduces the importance of trust, it also highlights the evolution between technology, trust and security from traditional commerce to electronic commerce. Information security policies are literally an all-in-one security policy resource with templates, advice, and instructions to help the business to generate practical, clear, and compelling information security policies. The study has also focused on the rapid development of EDI has posed a threat to the Indian automobile industries or has presented an opportunity to improve the efficiency and effectiveness of their services.

KEYWORDS

EDI, Productivity, Security, Templates, Automobile, Globalization etc.

INTRODUCTION

"EDI will change our lives, just as computers did. It will redefine the ways we work as it pushes us towards a knowledge-based society in which we pursue intellectual challenges while routine, noncreative tasks are assigned to computers" **Gene A. Nelson.**

Information security is considered as one of the most critical concerns in today's competitive digital economy. In today's fast moving market place, the business climate is changing at an unparalleled pace. The climate demands that organizations must adopt a new model for improving profitability, increasing competitiveness, allowing for globalization, shortening product life cycles, providing superior customer care and increasing shareholders value. Electronic Data Interchange (EDI) is a fundamental and permanent change; in the way, we do business. A strategic tool reduces expenses, streamlines business procedures, and creates a competitive advantage. The current related work done to support this service is first discussed, and then a new and simple, but effective protocol to achieve non-repudiation of receipt on a distributed and heterogeneous system platform is proposed.

Internet is now widely expected to function as part of the business and social infrastructure of the next century. Transactions conducted through the Internet will have enormous implications over the next few years for the International competitiveness of every country. It is a fundamental and permanent change in the way we do business. It is not only a strategic tool that reduces expenses and streamlines business procedures but also creates a competitive advantage. Global competitiveness, technological growth, opening of economy, expansion of markets and treating the whole world as a single market have necessitated the need for a continues flow of information as information works like lubricating oil to make economic system work better. Global competitiveness, technological growth, opening of economy, expansion of markets, and treating the whole world as a single market have necessitated the need for a continues flow of information as information works like lubricating oil to make economic system work better. EDI is not a technology, rather it is a solution to business needs i.e. paperless trading. Many companies are finding this as an innovative ways of incorporating it into other areas of their business and accross their entire trading channel. It has not only useful in purchase order and invoices, but is used to submit insurance claims and to apply for bank loans. Electronic Commerce, with Electronic Data Interchange (EDI) as its core, will be the most important element of the information infrastructure that will meet the business needs of a paper less world. EDI presents auditors with an excellent opportunity to employ continuous auditing techniques, on-line real time auditing tools and, potentially, new approaches to computerized techniques.

Figure-1

Sources: Author Compilation

[1]Professor and Dean (Administration), Rourkela Institute of Management Studies, Odisha, India, dr_kkpatra@yahoo.co.in

OBJECTIVES OF STUDY

- To integrate EDI with the current system security management.
- To highlight the risks and providing valuable guidelines on security models.
- To test the impact of rapid development of EDI is a threat or an opportunity.

REVIEW OF LITERATURE

Kavan (1991) EDI concerns multiple organizations-it brings buyer and seller together. As such, EDI adoption is a bilateral decision process. Adoption of the technology without corresponding adoption by other trading partner organizations renders the technology useless from an inter-organizational perspective.

Farbey (1995) the stage of EDI diffusion interpreted for each company was then further aligned with the approaches to EDI/IS planning. This indicated whether EDI was perceived as an incremental development, or part of a more enterprise-side change programme, within each organization and its supply chain networks.

Trust is defined as the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party (Mayer, Davis and Schoorman, 712:1995).

Some wisdom about Information Security

"Information Security is a combination of various factors. It involves technology, people and policy."

Sameer Kapoor, Executive Director, PricewaterhouseCoopers Private Limited

"Information Security is not just a technology issue this is a people and process issue too. The answer to this is education and awareness. You should talk to your employees."

Captain Raghu Raman, Practice Head, Special Services Group, Mahindra Consulting.

"Security has to move away from being a technology issue and become a business related issue."

Sunil Chandiramani, Partner, Ernest & Young

"There is a risk aspect to security too. Security breaches create a risk for the enterprise. So it's not just about hardware and software solutions."

Alok Shende, Industry Manager (IT Practice), Frost & Sullivan

"Security is now essential since it has become a business enabler. Enterprise Security should involve employees at all levels, customers and all entities that deal with the organization."

Anil Menon, Sr. Vice President, Operations, SecureSynergy

MANAGEMENT OF EDI PROCESS

EDI provides an innovative collaboration among business partners for fast, accurate, and secure exchange of business documents regardless of data format or network connectivity standards. The management of EDI process is explained below:

Figure-2

Sources: Author Compilation

EDI provides a greater degree of accuracy than traditonal paper equvalents. Adequate controlarenecessary for EDI environment including the applicaton software, interface software, translation software and communication software areas. It is not a technology, rather it is a solution to business needs. It is the combination of techniques and technology which facilitate paperless trading. It can be a cost saving tool or a cost burdern. It all depdnds upon how well the technology is implemented. Provision of adequate internal control is essential to ensure and verify the validity and timeliness of business transactions so it can be concluded that EDI is not just a new way to transmit data rather it is a new way to do business which adds complexity and sophistication to computerized accounting systems.

The adoption of EDI is changing how people think, changing busines procedures and changing how organization interacts with one another. It has great impact on accountants and auditors, and thus, these two professions must be aware of the strength and potential of EDI in changing their role. The main challenge are not technical, although they are important. The most vital tasks are preparation, planning and the concentration on internal systems. EDI, for all the pioneering work that is being done, is still in its infancy. That is why it is said that EDI is not a technology, yet it is a solution to business needs. EDI is not an end in itself rather, it is a tool that many companies in the world have used to restructure their internal operation and the way in which they relate to customers, and suppliers and authoritiies.

There is no shortage of testimonials as to the worthiness of EDI and its tangible benefits. Companies already using EDI can use the internet to transact EDI business with companies that have not yet invested in EDI softwares. Accountants are information officers, not just in accounting but also in the technology used to collect and process business information. Yet few are well versed in one technology with side ranging implications. Management accountant must be able to utilize the information technology to collect and process data, and distribute custom results to management.

As information technology changes, so do the underlying business operations and so must accounting services. Recent developments in electronic data interchange illustrate how changing information technology affects both business operations and management accounting practices. EDI is an online communication network what links the computer system of a vendor and its customer, reducing the need for conventional paper based communications.

A customer can place an order through the network directly to a vendor's system, avoiding the requisite paper work and time involved in traditional postal delivery. The vendor acknowledges the order, issues a delivery notice, and dispatches an invoice to the customer through the same network by mutual agreement, a manufacturer can even use the network to enter its customer's inventory status file in order to replenish inventory automatically. As management accountants are responsible for providing information to support executive decision-making, they are the organization best-qualified professionals to identify management's information needs. Their expertise is also indispensable in developing an EDI system that is integrated with the company's electronic data processing system. In fact, achieving an integrated EDI -EDP system is an ultimate goal of implementing EDI.

Thus management accountant often assume a leadership role here. While electronic data interchange proves opportunities for management accountants, it also offers them a challenge. Accountants must assert themselves as information officers, who understand not only accounting but information technology. Unfortunately, many accountants fall short According to several EDI specialists, many corporate accountants shy away from the technology. Some were dragged into the EDI development team. Others voiced concerns about possible problems with internal control. However, I believe the underlying problem is that many accountants know very little about EDI and simply dodge the issue. Corporate accountants must look more favourably on new technology and attempt to better understand it through continuous learning and professional development.

WHAT EDI IS NOT?

EDI is defined as the direct transfer of business information between computer system in different organization, (without human intervention) using widely agreed standards to structure the transaction or message data:

1. **Facsimile (FAX)**

   Documents transferred to a trading partner with FAX are not EDI because it requires a person to interpret the written data.

2. **Electronic Mail (E-mail)**

   This system eliminates the paper associated with FAX, but still it is not EDI because it requires a person to interpret and transfers the documents from one computer to another.

3. **Dedicated Computer Terminals**

   Any particular computer connected with the trading partner fully dedicated is not EDI because it requires to manually reeking internal data for transmission to a trading partner.

4. **Proprietary EDI**

   When data are exchanged electronically with a single trading partner using non-standard data formats called proprietary EDI. However, it partially satisfies the functions of EDI, but cannot be extended to business with other customers.

EDI COSTS

Table-1

Implementation Cost	Ongoing Cost
• Transaction software • Transaction set mapping • Communications software • Interface software development • Message transportation • Hardware acquisition • Security hardware/software • Project plan development • Training and education • Standard documentation	• Trading partner addition • Interface and application software modifications • Message transportation • Acquisition of new standard documentation • Maintaining EDI expertise • Legal negotiations • Implementation guide maintenance.

Sources: Authors Compilation

BENEFITS OF EDI IN TRADE FACILITATION

Table-2

EDI system leads to	Direct Benefits	Indirect Benefits
• Cost Saving, Materials savings i.e., stationery, envelopes, • Time Saving, • Physical savings i.e., postage, courier etc, • Reduced document handling, • Reduced re-keying of information.	• Increased productivity, • Increased customer services, • Increased cost savings, • Improved tracking of imports / exports, • Improved security.	• Improved National Efficiency, • Reduced costs to industry, government and nation • Improved competitive position, • Close relationship with trading partners, • Ability to broaden trading horizons.

Sources: Authors Compilation

To validate a trade debt, an organization's accounts payable department must match up three supporting documents: purchase order, receiving report, and vendor invoice. Because the first two documents are internally generated, they can be obtained through the internal EDP system. Vendors transmit their invoices through an EDI network. Thus, the organization must integrate its EDI system with its internal EDP system so that it can access and match the vendor invoices to purchasing orders and receiving reports. At the same time, the supplier's accountants receivable must be verified with a customer order and shipping report. The customer order is transmitted through the EDI network. With proper matching or the customer order to an internally generated shipping report, the supplier's accounts receivable function can produce an invoice and transmit it back through the network to the customer.

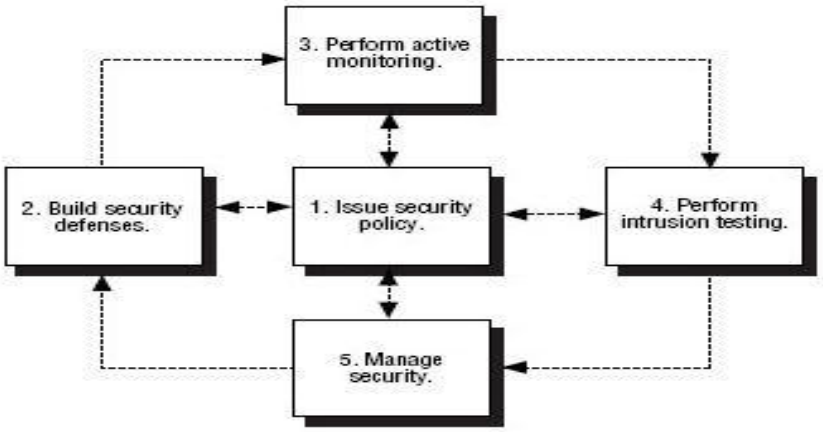
With minimal human interference, an integrated EDI-EDP system can update both customer and vendor accounts and record necessary journal entries in real time. Efficient EDI system requires prompt payments and receipts. Rather than send cheque payments through the mail, EDI advocates recommend that payments be routed through the bank's electronic funds transfer (EFT) system. If banks do not participate in an EDI network, trading partners can still instruct their banks to transfer funds over the telephone to the vendor's account. However, without tangible evidence of payment authorization, telephone communications pose an internal control problem. In an EDI network, the bank's EFT network is virtually integrated with the customer's EDI system. Trader's requests for the payments transmitted through an EDI network to their banks are well documented in electronic form. EDI communication improves other business operations. Establishing EDI communication allows shippers to monitor the progress of their customer deliveries, particularly important for companies with large geographical markets.

Electronic Data Interchange (EDI) is a computer-to-computer or application-to-application exchange of business information in a standard format. In 1992, there were over 31,000 known EDI users, with a steady increase since 1987. EDI users can be found in such industries as transportation, retail, grocery, automobiles, warehousing, pharmaceuticals, healthcare and financial institutions. As a comprehensive book on EDI, several parts of the book deal more with the operation and setup of such a network. This leads into the areas that explain in technical detail the security and auditing of EDI networks. Beginning with the basics of EDI, the book walks through the pros and cons of such networks. It gives guidelines for who should implement and use it, operating issues, risks, control concerns and more. These sections are brief and to the point, suitable to give to non-technical managers who may be considering EDI as a solution.

IT SECURITY GOVERNANCE

Information security governance is a critical part of an organization's overall enterprise governance. The board and executive management are to ensure that there is an effective program of activities that provide continued and adequate assurance. Because of these considerations, it is a necessity to vigorously continue the information security journey. Information security managers can consider and act on each of the aspects of information security governance.

Figure-3: IT Security Governance



Sources: ISACA, CISM Review Manual 2005

HOW EDI IMPROVES ACCOUNTING SYSTEMS

This is an age of transition. Computers and systems concepts are not completely used by both small and business. Hence, accountants should be able to maintain books of accounts in computer-free environments. The electronic computer does not affect accounting principles because they are independent of processing methods, but the enhanced processing capabilities of the computer have improved the ability of organizations to implement advanced management information systems, which were not practical using manual methods. Therefore, those involved in accounting and information system of an organization should understand the capabilities of computers and the changes in data processing because of the computer. In organizations using the computers for data processing, the accountants should be able to understand computer data processing methods in organization and processing of computer files. The ability to do analysis to provide meaningful reports and/or to prepare statements depends, largely on the existence of appropriate files and processing capabilities. The accountants should also understand quality control for computer processing because the reliance, which should be placed on computer-based records, is dependent on the quality of files and accuracy of processing. Hence, it is appropriate for any programmed on accountancy to have systems curse prior to accounting and auditing courses. For an accountant, background in computer programming is also desirable though not essential; however, an accountant must be able to use front-end tools.

Benefits of EDI systems

There are two types of benefits generally available in EDI system:

Direct Benefits

- EDI system leads to cost and time savings
- EDI system reduced errors and increased productivity
- EDI system increased customer services and cost savings
- EDI system improved Security.

Indirect benefits

- EDI system improve National efficiency
- EDI system reduces costs to industry and government
- EDI system improves competitive position.
- EDI system develops closer relationship with trading partner.

PARTIES INVOLVED IN INFORMATION SECURITY

The Board of Directors

The Board of Directors manages the business of the company in so far as it has not delegated management of the company responsibility among other things for the upper management of the company and the issuing of necessary instructions, as well as «the supervision of persons entrusted with management of the company, especially with regard to the observance of laws, statutes, regulations and instructions. The committee believes that it is important for corporate Boards to be fully aware of the risks facing the business and that it is important for shareholders to know about the process by which companies manage their business risks.

The Management

Company management manages the business of the company together with the board of directors. At the same time, every person appointed to company management is subject to the organizational rules, which define the necessary positions for the management of the company. In practice, it is difficult to prove a violation of this principle by company management or even an individual member of company management in a court of law because in the absence of conflicts of interest, the legislator always acts under the assumption of careful and loyal company management.

The Public

The public is relevant as far as it is made up of potential clients, investors, stockholders, employees and business partners. The public also includes specific interest groups or organizations, which feel that they are affected in any way by the activities of the bank. As a rule, the individuals who make up the public are poorly organized, however in particular situations they can count on the support of the media or specific interest groups or organizations.

The Clients

Clients are personally and strongly affected by deficiencies in the information security of their bank. They are however poorly organized and in the event that damages are determined, can, if at all, only assert their own claims within the limits of the law at great expense. In addition, as non-expert outsiders, they often do not assess the information security provided by the bank well and are subject to the danger of moral hazard.

The Stockholders

Stockholders are partners in the corporation and have residual claims to the earned profits. Stockholders specialize in the assumption of business risks and entrust the management of the business to a management team. The separation of these two important functions (residual claim and the right to coordinate) theoretically makes it possible to engage the best-suited people in management and tends to lead to a more risk-accepting and dynamic company because business risks can be distributed widely.

Possessors and Owners of Information

It includes lawyers differentiate between possession and ownership. A Possessor is any person who holds custody of property of his own or another person. If property entrusted to someone's care is damaged, then the rightful owner should bear the damages. However, the owner can assert a claim for the compensation of damages if misconduct on the part of the possessor is proven. In connection with this, it is advantageous when the conditions and expectations of the owner are disclosed at the transfer of possession, and accepted and confirmed by the recipient. This model also has a wide range of applications in the field of information security in the definition of rights and duties in the handling of information.

INDIAN AUTOMOBILE INDUSTRIES: AN OVERVIEW

The automotive sector is one of the core industries of the Indian economy, whose prospect is reflective of the economic resilience of the country. Continuous economic liberalization over the years by the government of India has resulted in making India as one of the prime business destination for many global automotive players. The automotive sector in India is growing at around 18 per cent per annum. "The auto industry is just a multiplier, a driver for employment, for investment, for technology" The Indian automotive industry started its new journey from 1991 with delicensing of the sector and subsequent opening up for 100 per cent FDI through automatic route. Since then almost all the global majors have set up their facilities in India taking the production of vehicle from 2 million in 1991 to 9.7 million in 2006 (nearly 7 per cent of global automobiles production and 2.4 per cent of four wheeler production).

The cumulative annual growth rate of production of the automotive industry from the year 2000-2001 to 2005-2006 was 17 per cent. The cumulative annual growth rate of exports during the period 2000-01 to 2005-06 was 32.92 per cent. The production of the automotive industry is expected to achieve a growth rate of over 20 per cent in 2006-07 and about 15 per cent in 2007-08. The export during the same period is expected to grow over 20 per cent. The automobile sector has been contributing its share to the shining economic performance of India in the recent years. With the Indian middle class earning higher per capita income, more people are ready to own private vehicles including cars and two-wheelers. Product movements and manned services have boosted in the sales of medium and sized commercial vehicles for passenger and goods transport.

Videsh Sanchar Nigam Limited (VSNL) was the first to offer EDI infrastructure in the country. VSNL installed the EDI system in Mumbai with access nodes at New Delhi, Kolkata and Chennai in 1993 and offered this service to potential users. In the year 1997, Chennai based Satyam Infoway (p) Limited, announced it is implementing a pilot project for the Indian Automobile industries represented by the Automotive Components Manufacturer's Association and Association of Indian Automobile Manufacturer has titled "Auto Pilot". The project consists of network automobile manufacturer's components suppliers and dealers with an EDI application and speed up the transaction cycle in the industry. 11 companies such as Bajaj Auto, Telco, Mahindra and Mahindra, Ashok Leyland, Mico, Shriram Pistons, Sundaram Clayton, Sundaram Fasteners, Lucas TVs, Remson Industries, Brakes India etc covered the project. Other wide area network (WAN) services providers such as Mahindra Network Service Limited (MSNL); global Telecom Service Limited (GTSL) besides VSNL also participated in the project. A connection

between the three service providers has already been established through a special gateway provided by VSNL. A comparison of four organizational models for supply chain coordination is given below:

Table-3: Comparison of Four Organizational Models for Supply Chain Coordination

Characteristic	Regional Clusters as Auto-coordinators for the Sourcing Network	Sourcing Networks Without Local Links	International Partnerships Linked to Global Market	Sourcing Networks Coordinated by Traders
Profile of SC Coordinator	Largest and more entrepreneurial manufacturing firms in the cluster	OEMs serving as first tier suppliers of TNCs, or dominant firms belonging to the SC	Joint ventures or partnerships between local and international OEMs, both integrants of SC	Local or international service and knowledge firms: traders, brokers, neutral third parties
Production and Logistics Activities	Performed by producers in the cluster, assistance from governmental agencies and industrial groups	Performed by the SC coordinator, logistics possibly performed by 3PLs.	Performed by the SC coordinator, logistics possibly performed by 3PLs.	Production supervised by the trader, logistics could be performed by the coordinator or by 3PLs directly managed by the coordinator
Participation of Domestic Producers	Intense and direct	Indirect and limited	Direct but limited	Intense and direct
Supplier's Control	No control, voluntary participation to be a part of the strategic network	Tight control, suppliers owned by the 'hub firms' playing the coordinator role	Tight control, suppliers owned by the coordinator or forming an alliance or joint venture	No control, voluntary participation to be a part of the strategic network
Supplier's Upgrading	Ineffective and with high differentials, mainly in conformance to international quality standards	In product and process, but no participation in other high-value added activities like R&D and commercialization	In product and process, but no participation in high=value added activities like R&D and commercialization	In product and process, partial development of design capabilities, opportunity to be a specialized producer
Contribution at Regional Level	Sustainability and competitiveness of existing, SMEs, more export activity.	Employment, human resource upgrading, industrialization and export	Employment, human resource upgrading, industrialization and export	Sustainability and competitiveness of existing SMEs, more export activity
Types of Network	Horizontal integration supported by social networks	Vertical integration, mainly arm-length ties	Vertical integration, mainly arm-length ties	Horizontal integration, partnership ties
Coordinator's Effectiveness	Ineffective, limited capabilities, (financial, logistics, establishment or forward linkages)	Effective, high production expertise, with financial resources and forward linkages	Effective, high production expertise, with financial resources and both backward and forward linkages	Effective, high networking capabilities, profound knowledge about local producers capabilities, customer and service-oriented

Sources: The ICFAI University Journal of Supply Chain Management, June 2009

IT-ITeS Industry: Growth Performance

The Indian Information Technology-Information Technology-enabled Services (IT-ITeS) industry has shown remarkable resilience in the year 2007-2008. Continuing on its established record of accomplishment, the overall Indian IT-ITeS revenue aggregate is estimated to have grown by over 33 percent to reach US$ 64 billion in fiscal year FY 2007-2008 as compared to US$ 48.1 billion in FY 2006-07.

Table-4: India Domestic IT/ITeS Market Size (in Rs. Cr.)

Market Segment	2007	2008	2009	Growth 2008 Over 2007	Growth 2009 over 2008
Software	7,823	9,628	11,300	23.1%	17.4%
Services	20,920	25,092	29,934	19.9%	19.3%
Hardware + Others	52,890	59,465	63,703	12.4%	71%
Total Domestic IT Market	81,633	94,185	1,04,937	15.4%	11.4%
Total Domestic ITeS Market	4,468	6,846	9,637	53.2%	40.8%
Total Domestic IT/ITeS Market	86,101	1,01,031	1,14,574	17.3%	13.4%

Sources: The Economic Times, March, 2009 p.2

Recently, Cyber laws bill has been cleared by the cabinet and passed by the Parliament. It is now in force in the country. In India, with its high emphasis on indigenization with a consequently large number of suppliers, EDI should prove viable and most useful in India. EDI presents audit professionals some novel challenges. EDI thus demands innovative thinking and exercise of judgment. An auditor should not shrink from these challenges. EDI project, if successfully implemented, would save costs and bring about standardization in the Indian automobile industries.

Indian automobile industry has come a long way to from the era of the Ambassador car to Maruti 800 to latest M&M Xylo. An industry is highly competitive with a number of global and Indian companies present today. It is growing at a pace of around 18% per annum for the last five years and is projected to be the third largest auto industry by 2030 and just behind to US & China, according to a report. The industry is estimated to be a US$ 34 billion industry.

Indian Automobile industry can be divided into three segments i.e. two wheeler, three wheeler & four-wheeler segment. Two wheeler segments enjoys 75% market share of automobile industry, followed by passenger vehicles with the 16% share of market. Three wheeler segments have merely 4% share in domestic market. Indian dominates the domestic two-wheeler market as well as foreign players such as Hero Honda, Bajaj Auto, Honda Motors, TVS Motors, and Suzuki etc. Maruti Udyog and Tata Motors are the leading passenger car manufacturers in the country. In addition, Suzuki, Yamaha, etc consider India as strategic market. The major players have not left any stone unturned to be global. Major of the players have got into the merger activities with their foreign counterparts. Like Maruti with Suzuki, Hero with Honda, Tata with Fiat and latest Mahindra with Renault.

EDI SECURITY, CONTROL AND AUDIT

EDI is an immensely valuable technology. It promotes efficiency and productivity. Companies implementing EDI can reap tremendous cost savings, respond more quickly to changes, in the market and realize critical strategic objectives. EDI can improve the effectiveness of the audit process by giving auditors more reliable information, and quick access to it than in a paper environment. An auditor must be conversant with the computer and EDI technical ability in order to evaluate application and general controls to substantiate evidential matter.

With the rapid increase in the number of companies in the present century, this profession has assumed as an ever-increasing role in the business community. It has a wide usage and implies a thorough examination of the books of accounts to repose confidence on shareholders in particular and other users of financial statements in general. An auditor played a vital role in instilling confidence in the public at large with regard to company form of organization by revealing the facts to them. This will further enhance the reliability of the audit report and will help in continuing to repose the confidence of end users of the accounting services.

Whether the business is using traditional paper documents or implement EDI, adequate internal controls are essential to ensure and verify the validity and timeliness of business transactions. It would definitely useful to an auditor for the check and balances in the books of accounts. In a traditional business, verification can be made with paper documents, for example, a purchase order is used to ensure and verify that an order was placed. To an auditor, paper documents such as purchase orders serve as the major form of a transactions existence. However, under EDI audit, the auditor loses paper evidence; but as the EDI, system provides evidence in electronically and given more details than the practical paper. It is evidenced electronic copy of the EDI documents can make the audit process more efficient and effective than with additional paper documents. Retrieval facilities are faster in EDI system comparing to the file cabinet which all information available in scattered way.

It is very much difficult for an auditor to check the referred documents, which are available in electronic form. Electronic documents lack tangible and it requires having external qualities, which helps to substantiate the authenticity. EDI is an immensely valuable technology. It promotes efficiency and productivity. Companies, implementing EDI can reap tremendous cost savings, respons more quickly to changes in the market, and realize critical streategic objectives. EDI is an on-line communication network which links the computer system of a vendor and its customer in a most efficient way.

CONCLUSIONS

EDI is not a technology, yet it is a solution to business needs. EDI is not an end in itself rather, it is a tool that many companies in the world have used to restructure their internal operation and the way in which they relate to customers, and suppliers and authorities. The adoption of EDI is changing how people think, changing busines procedures and changing how organization interacts with one another. It has great impact on accountants and auditors, and thus, these two professions must be aware of the strength and potential of EDI in changing their role. There is no shortage of testimonials as to the worthiness of EDI and its tangible nenefits. The most vital taks are preparation, planning and the concentration on internal systems. EDI, for all the opioneering work that is being done, is still in its infancy. Companies already using EDI can use the internet to transact EDI business with companies that have not yet invested in EDI softwares. To strengthen the findings of the paper, it has been supported by a survey carried out by the Institute of Chartered Accounts of India and the automobile industries in India.

REFERENCES

1. Anderson, Shannon W., & Laner, William N. (2002, October). Using Electronic Data Interchange (EDI) to improve the efficiency of accounting transactions. ***The Accounting Review***, 77(4), 703-729.

2. (1998, June 22). Doing business in the internet age. ***Business Week,*** 121-172.

3. Cash, J., & Konsynski, B. (1985, March-April). Is redraws competitive boundaries. ***Harvard Business Review***, 63(2), 134-142.

4. Farbey, B., Land, L., & Targett, D. (1995). A taxonomy of information systems applications: The benefits evaluation ladder. ***European Journal of Information Systems***, 4(1), 41-50.

5. Gordon, Judith R. & Gordon, Steven R. (1999). ***Information Systems-A Management Approach***. Harcourt Bracae College Publishers: The Dryden Press.

6. Helms, Glenn L. & Mancino, Jane. (1998, April). The Electronic Auditor. ***Journal of Accountancy,*** 23-25.

7. Hwang, K. T., Pegels, C. C., Rao, R., & Sethi, V. (1993, December). Electronic data interchange - state of the art. ***Journal of Systems Management,*** 44(12), 12-15.

8. Janardan, N. (2000). ***Electronic Commerce***. New Delhi: Indian Institute of Foreign Trade, 26-46.

9. Janardhan, N. (1998). ***Why is Electronic Commerce a Strategic Policy Issue for India***. Jaipur: Electronic Commerce, Data and Expo India (P) Limited.

10. Johnson, J. Z. (1998). ***Network Security Programs: Process and Metrics for the Real-World***. White paper, Internet Security, Systems, Inc.

11. Kavan, C. (1991). ***The adoption of inter-organizational systems: The example of electronic data interchange*** (Doctoral dissertation). University of Georgia.

12. Khan, T. A. (1998). ***EDI/EC in the Asian Region.*** Jaipur: Electronic Commerce, Data and Expo India (P) Limited.

13. David, Kosiur. (1997). ***Understanding Electronic Commerce***. Washington: Microsoft Press.

14. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. ***Academy of Management Review***, 20(3), 709-734.

15. Power, R. (2001). ***Computer Security: Issues and Trends.*** 2001 CSI/FBI Computer Crime and Security Survey, VII (1).

16. Shapiro, C., & Varian, H. (1999). ***Information Rules - A Strategic Guide to the Network Economy.*** Boston, Massachusetts: Harvard Business School Press.

17. Venkatraman, N. (1994, winter). IT-enabled business transformation from automation to business scope redefinition. ***Sloan Management Review,*** 73-88.

*****

NOBLE APPROACH TOWARDS EFFECTIVE ELECTRONIC SERVICES DELIVERY IN TRIBAL SECTOR

A. K. Hota[2] Ashis Kumar Mahapatra[3] Bikram Kesari Rath[4]

ABSTRACT

In a multi-application environment, applications multiply very fast and inter-communication complexity grows at an alarming rate. Simplifying the inter-application communication is a long desired goal for several reasons. Specifically, in e-Governance type of solutions where multiple applications are implemented on different platform and governed by various inter dependent departments.

The inter-application communication is of paramount importance for seamless service to citizens. Inter Operability between different departments needs Inter Operable Services and as the service consists of many small processes. Therefore, Inter Operable Service is again achievable by only Inter Operability between different heterogeneous processes. It demands multiple applications calling multiple services creating a confusing mesh of service calls. Fail-over, load balancing, etc., soon become nightmares and Service management is challenged. Distributed services and even more distributed service consumers make an inefficient architecture – resulting in lower citizen value from systems.

The objective is to develop a prototype "Integrated Data Engine" that centralizes inter-application service calls as a proxy, which provides fault tolerance, security, and load balancing capabilities etc with focus in tribal development sector.

KEYWORDS

e-Governance, Tribal Sector, SOA, e-Government, Web Services etc.

INTRODUCTION

e-Governance cannot be introduced in the whole country across government organizations at one go. e-Governance is an integral part of reforms in governance and each organization needs to embed e-Governance systems within the organization in a seamless way. However, different organizations are not, presently, at the same level of e-preparedness. There has to be a step-wise approach to e-Governance so that outcomes are maximized and citizens reap early benefits from e-Governance. These steps are mentioned below:

1. **E-Preparedness**

   A certain level of preparedness is essential for any e-Governance project, in the form of existence of basic infrastructure and human resource capabilities of the organizations. A clear assessment of actual e-preparedness of an organization should be first carried out while conceptualizing any project. Efforts should be made to enhance the e-preparedness to the desired levels.

2. **Identification of e-Governance Projects by each Organization / Entity**

   Each organization would have to identify areas/activities falling under its functional domain, which could benefit from e-Governance. This identification has to be based on the needs of the citizens.

3. **Prioritization**

   Once the e-Governance projects to be undertaken by an organization have been identified, they would require to be prioritized. This should be based on simplicity of the project, ease in implementation and benefits to the citizens. However, capacity building for the more complex projects should remain in focus. Simple projects serve to bring out the potential benefits of e-Governance. These create a demand for more such initiatives. However, the more complex initiatives have the potential of bringing the larger benefits of good governance to citizens.

4. **Business Process re-Engineering**

   Whether it is for providing information and services to the citizens or for streamlining the internal functioning of government organizations, each e-Governance initiative would have to be accompanied by a systematic analysis of the governmental processes involved and tested on the anvil of simplicity and desirability. This would lead to redesign of processes using technology. The process would result in, if required, changes in forms, processes, structures and laws

[2] Senior Technical Director, National Informatics Centre (NIC), Department of Electronics & IT (DeitY), Ministry of Communication & IT, Government of India, Odisha State Centre, Odisha, India, ak.hota@nic.in
[3]Scientist-D, National Informatics Centre (NIC), Department of Electronics & IT (DeitY), Ministry of Communication & IT, Government of India, Odisha State Centre, Odisha, India, ashis.mahapatra@nic.in
[4]Reader, Utkal University, Odisha, India, vkramus@yahoo.com

and regulations. The exercise should centre on the needs of the citizens. Business process re-engineering in government organizations is a complex task because, as opposed to private organizations, governmental processes and structures are designed or regulated by various statutes, rules, regulations, instructions etc. changing them would require a complete understanding of the functioning of individual government organizations and laws and regulations associated with them on the one hand and technological applications and the needs of the citizens on the other. Thus, this exercise would form the backbone of e-Governance initiatives.

5. **Developing Technological Solutions**

Every e-Governance initiative would require its own technological solution. However, there would be commonalities across Union, State and local government levels. Further, there would be need for sharing of information and establishing connectivity across organizations at different levels. This would require standardization of basic requirements, adoption of interoperable platforms and creation of data storage and retrieval systems. In the end, the technological solution would have to be modified according to the specific needs of the organization with the help of field experts. However, care has to be taken to ensure that the country does not follow the beaten path in adopting technological solutions. Due to rapid strides in the development of Information and communications Technology, there is a strong case for 'leap-frogging' in the selection of technology and applications to achieve better results. Applications mounted on mobile telephone sets and other hand-held devices are examples of such advancements. The technological solution should be able to provide a simple interface to the citizens, be cost-effective, promote efficiency, be sustainable and reliable and lend itself to scalability.

6. **Implementation of e-Governance Projects**

Only when the above-mentioned steps have been undertaken should e-Governance projects be implemented. This may be in the form of a pilot project initially, testing the technological solution and the procedural and functional inputs. However, the business process re-engineering should be such that once the e-Governance project has been successfully tested on a pilot basis, the scaled up system should be able to make e-Governance irreversible, i.e. it should not relapse into the pre-e-Governance mode. In the past, there have been numerous instances of successful projects not being scaled up and of lessons not being learnt from failed projects. Thus, each pilot project should invariably be evaluated and the lessons learnt including why it did not lead to scaling up. Failed projects should be redesigned to remove shortcomings. This would ensure that organizations do not become a graveyard of pilot projects.

E-GOVERNANCE IN TRIBAL SECTOR

e-Governance has to be implemented across different departments and organizations with a wide spectrum of activities and with varying levels of readiness for e-Governance. Achieving the desired results would, therefore, require the fullest political backing, a determined and resolute approach by all organizations and departments of Government as well as active and constructive participation by the public. It would require providing institutional and physical infrastructure for taking e-Governance initiatives across our cultural and regional diversities; more importantly, it would require the creation of an environment that would encourage the adoption of ICT. Thus, apart from the technical requirement, success of e-Governance initiatives would depend on capacity building and creating awareness within government and outside it.

The National e-Governance Plan (NeGP), takes a holistic view of e-Governance initiatives across the country, integrating them into a collective vision, a shared cause. Around this idea, a massive countrywide infrastructure reaching down to the remotest of villages is evolving, and large-scale digitization of records is taking place to enable easy, reliable access over the internet. The ultimate objective is to bring public services closer home to citizens, as articulated in the Vision Statement of NeGP.

Tribal portion of e-Governance portfolio consists of various citizen facing solutions specifically for tribal requirements but then these solutions create value when they are integrated with other solutions that interface with population in general. Therefore, for example, a tribal scholarship scheme might create extra value if interfaced with a general population solution like registration of birth. This integration is the challenge addressed in the work here. [**1**] [**2**]

TRIBAL SECTOR AS FOCUS OF STUDY

Tribal communities are vulnerable because not only they are poor, asset less and illiterate compared to the general population; often their distinct vulnerability arises from their inability to negotiate. The following persistent problems have largely remained unattended:

- Land alienation,
- Indebtedness,
- Non-implementation of the Forest Rights Act,
- Lack of proper rehabilitation,
- Shifting cultivation,
- Poor utilization of government funds,
- Poor delivery of government programmes.

These issues needing urgent attention are under the jurisdiction of department of:

- ST & SC Development,
- Environment & Forests,
- Rural Development,
- Panchayati Raj,
- School & Mass Education,
- Higher Education,
- Housing & Urban development,
- Women & Child Development, etc.

A systemic change is needed in the way ST & SC Development Department function; the approach must change from simply spending the own budget through narrow departmental schemes to **knowledge-based advocacy** with other concerned departments. ST & SC Development Department must highlight the failure of governance that deprives the poor tribal's of access to elementary services, and put pressure on the concerned departments to ensure better policies and delivery in tribal regions.

When inter-system communication takes place we find [**3**] [**4**] [**5**]:

- Information given in one process does not match another process,
- Contradictory information seems to be present,
- Mis-allocation of resources (funds) might be taking place,
- Eligible candidates might be getting rejected,
- The question is how to integrate such separate "databases" to create a true picture in service-oriented architecture apps.

PROBLEM STATEMENT

The use of service-oriented (SO) distributed systems is increasing. Within service orientation, Web Services (WS) are the de facto standard for implementing service-oriented systems. The consumers of WS want to get uninterrupted and reliable service from the service providers. However, WS providers cannot always provide services in the expected level due to faults and failures in the system. As a result, the fault management of these systems is becoming crucial.

There are at least three factors that contribute to the performance or lack there off, in web services:

- The first is network transaction time. It means how long it takes the client to make a request to the remote Web Service.
- The second factor is the time to take to handle the message. Specifically the XML parsing, any flow management, invocation of service, and finally response encoding. Last performance factor is the service time itself takes to execute.
- This third factor is often the major culprit in ill-performing services. It is easy forgotten that useful code can take some time to perform its function [**6**].

OBJECTIVES OF STUDY

Major objective of this Paper is to Design a data engine that demonstrates the simplification of SOA implementation as inter-application communication grows and underlying service calls, security, and fault tolerance etc. become almost impossible to manage efficiently.

RESEARCH METHODOLOGY

A method has been proposed to centralize some of the various mechanisms for extra efficiency, ease of development, better security handling and robust fail-over in e-Governance solutions [**7**][**8**] with Tribal Development as focus of study.

To achieve the above objective following in depth study has been considered:

- Study of Government Processes at ST & SC Development Department like ITDA, MADA, MICRO, CLUSTURE, PTG, SCSTRTI, OSCSTFDC, ATLC, DWO Office, Educational Institutions.
- Study & Development of Databases for ST & SC Dev Department for Individual Family Survey, Forest Right Act (Oriya Language integration), Schemes Monitoring of Article -275(1), SCA to TSP, Self-Employment Training for ST & SC, Digital Preservation of Tribal Artefacts.
- Study of Databases of other departments associated with tribal Development like BPL Data (H & UD + RD Department), EPIC Data (Election), Farmer I Card System (F&CS Department), e-Scholarship (Post-Matric / Pre-Matric) (Mass Education / Higher Education Department), Land Record (Revenue).
- Defining Standardization of Meta Data for various Governance processes related to tribal development for Village Data, Block, District, Tehasil, BPL, Other Government schemes.
- Developing web services for various business processes.
- Developing Central Repository of Services.
- Identifying dependant service(s) of any composite services.
- Developing & Integrating Data Engine.

REVIEW OF LITERATURE

Table given below shows the summary of related works related to technical terms and management of distributed systems and management of web services. The works on management of distributed systems are related to fault management of WS as WS are distributed in the network

Table-1: Summary of Related Works

Reference Paper	Work and contribution	Limitations
D. Ardanga et al. [**9**]	Present classification of faults in web services and a self-healing platform for recovery of faults. This paper classifies WS faults into three levels: a) infrastructural and middleware level, b) web service level, and c) web application level. This work presents a self-healing platform, which implements the run time service oriented fault analysis and recovery actions.	This paper shows the theoretical representation of the system. It does not present any experimental data.
Weiping He [**10**]	Proposes an infrastructure to implement failure recovery capabilities in WS Management Systems. He classifies failures in WSMS into five categories -functional failures, operational failures, semantic failures, privacy failures, and security failures.	Does not show any experimental data.
Masoud Mansouri et al. [**11**]	Presents a functional model for monitoring of distributed systems. Presented a monitoring model that has a set of monitoring functions: a) generation of monitoring information, b) processing of information, c) dissemination of information, and d) presentation of information.	Presents monitoring model only. Does not include overall management solution. It also does not present any results.
A. Benharref et al. [**12**]	Presents a web service based architecture for detecting faults in web services.	Results were not shown. For deadlocks the system does not work.
Ertugrul Akbas [**13**]	Presents RMI and JAVA based architecture of system independent and distributed fault management system	This one also does not show any result.
Pankaj Kumar, HP [**14**]	Paper presents some best practices that can be used by web services during distributed system management	Theoretical presentation. No quantitative result shown.

Sources: Authors Compilation

All of the papers cited in the related works lack quantitative results. The authors did not show any experimental data. Therefore, we do not know how the systems perform. There are also some additional open questions regarding the proposed systems. 1. How long does it take to detect and manage faults? 2. Can it detect all the faults? SOA and Web services are state of the art and maturing fast. Literature is abundant about various challenges and issues in the areas. Web services are the standard architecture of IT implementations. E-Governance is fast becoming a large undertaking and must address the challenges identified in the research and existing implementations. One of the challenges is how to manage the explosive growth of services in any particular project implementation.

WEB SERVICE DESIGN

The services are designed to fulfill the condition that cross-application web calls are made where complex interchange of data takes places.

Web Service Calls

Reliability and ubiquity are critical in a typical multi-tier computing infrastructure that uses a pool of web applications and web services in tandem. However, a number of failures occur nonetheless because of bugs in source code, uncaught exceptions, thread deadlocking, network failure, data type mismatch, service failure etc. Most common faults and their corresponding remedies are detailed in below table 2.

Table-2: List of Common Faults and Fixes in Multi-Tier Architecture

Faults	Fixes
Source Code Bugs	Reboot Tier and / or service, send notification to administrator,
Uncaught Exceptions	Micro-boot service or component,
Deadlocked Threads	Micro-boot service, abort request,
Buffer Contention	Reparation the memory across various buffers,
Aging	Reboot to reclaim leaked resources,
Read / Write Contention on Table Block	Repartition table to balance accesses ground partitions,
Service Request Overload	Redirect to another replica service trough load balancer algorithm,
Data Type Mismatch	Convert to appropriate data type,
Service Source Code Malfunction	Find out the return result expect from the failed service and try to get the same result from the combination of other available services.

Sources: Authors Compilation

Self-healing service integration needs to deal with these kinds of failure, which can show up in both the system level as well as the logical level. Once the failure has been detected, the service needs to be able to make an optimal choice and solve the problem by reconfiguration. This self-healing behavior must adhere to specified system-level policies. We can either cancel the request to try again at another time hoping the problem fixes itself, or we can abandon the service and try to find another service that would process the request. Self-healing software also needs to be equipped with a good test harness and test data to ensure the system responds to stated faults while keeping system-level policies in mind. We could produce a set of assertions based on the application's semantics. These assertions can then be validated during run-time against a known set of data.

PROPOSED DATA ENGINE

Table-3: Data Engine - Recovery Action Plan

Sl. No.	Root Cause	Recovery Action	On Failure Follow
0001	Connection error	Retry to same service	0002
0002	Network Failure	Follow different path to the same service through PFA (Path Finder Algorithm)	0003
0003	Server Down	Find & reconnect to the backup server	0004
0004	Service Not Available	Try to get the same result from different composition of services.	
0005	Request Overload		

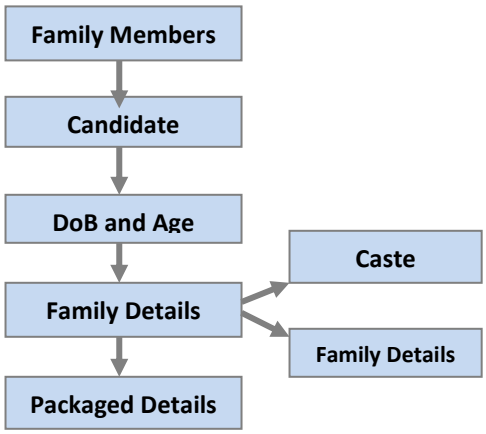
Sources: Authors Compilation

e-Governance space demands architectural elements like data-engine view of web services because the scale of implementation is too large. As intra and inter application integration takes place, a data engine view must be analyzed to understand the dynamics and challenges. It is a timely topic of research to explore the space of data engine view of web services taking one particular domain of e-Governance like that of Tribal space. Before services are designed and implemented, it is essential to set standardization practices in place. Process standardization will be crucial but not within the scope of this work. Data standardization is another important component of the exercise, which is included in the scope of current work. It is observed after extensive research in various e-governance databases and applications that various metadata need to be created for standardization. Within the scope of this thesis, one such important Meta data set is personal identification information. This information is utilized and referred in maximum number of applications.

SERVICE CALLS

The testing service calls structure is depicted in a block view below:

Figure-1



Sources: Authors Compilation

Each service is detailed below with well-commented code explaining:

- Text definition of service,
- Parameters IN,
- Parameters OUT,
- Error handling,
- Code snippet.

RESULTS FROM WEB-SERVICE CALLS

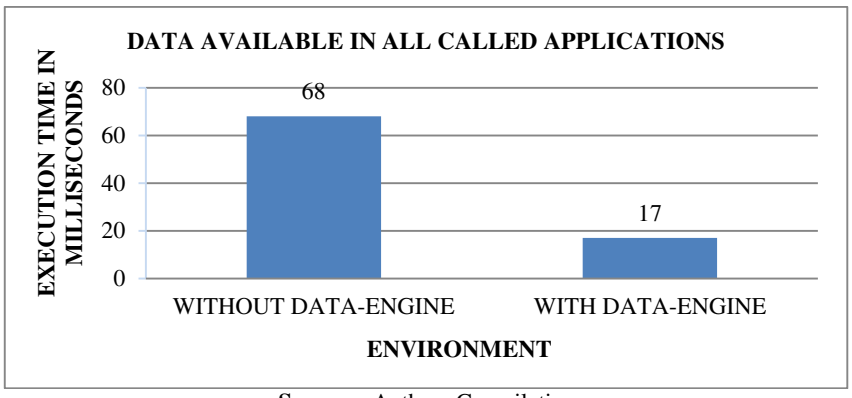
Using the experimental web-service tools and call structure, sufficient data is collected to give rise to results and analysis that is a good representative set of real world.

When services are called for a specific record, following scenarios are tested:

A) Scenario – When applicant's data and family data is found in all applications

- In this scenario, data for a given applicant is found in DOB dataset, and in income BPL dataset and in the Land Records dataset.

Graph-1



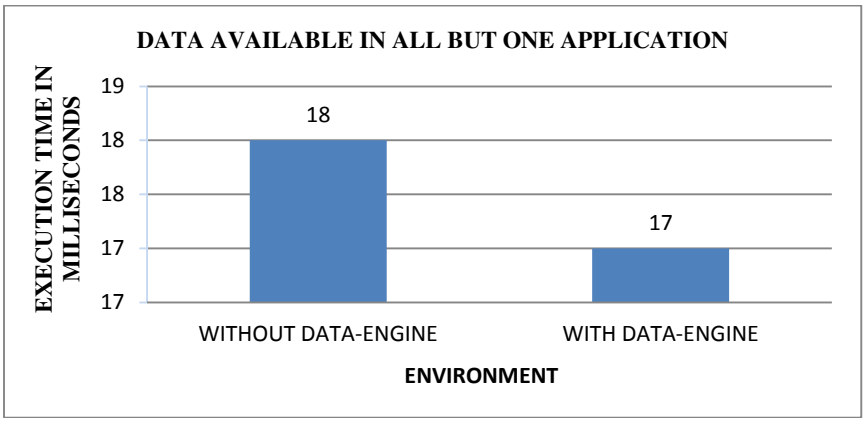
Sources: Authors Compilation

- Calling Services without data-engine takes 68 milliseconds while calling with data-engine view takes 17 milliseconds. That results in an improvement of 51 milliseconds over 68 mill seconds, which is an improvement of 75%.

B) Scenario – When applicant's data and family data is found in all but one application

- In this scenario, data for a given applicant is found in all the datasets applications but not in Income dataset.

Graph-2



Sources: Authors Compilation

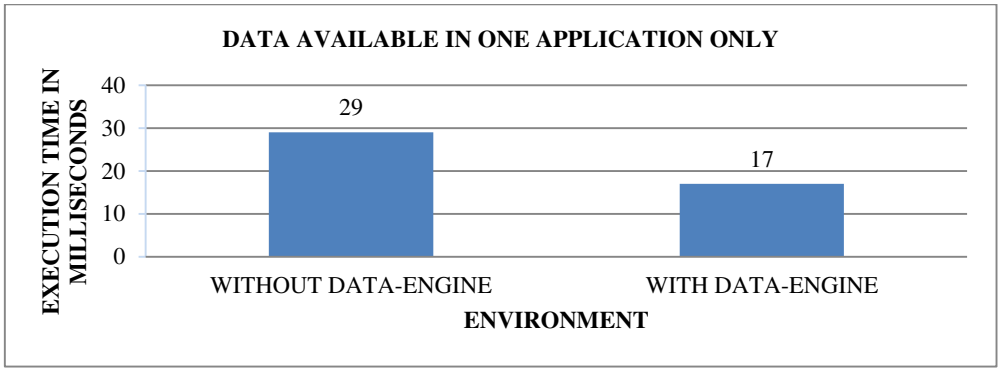
- Calling Services without data-engine takes 18 milliseconds while calling with data-engine view takes 17 milliseconds. That results in an improvement of one millisecond over 18 milliseconds, which is an improvement of 5%.

C) Scenario – When applicant's data and family data is found in one or two databases but is absent in other applications

- In this scenario, data for a given applicant is found only in one application and is absent in all other datasets.

-

Graph-3



Sources: Authors Compilation

- Calling Services without data-engine takes 29 milliseconds while calling with data-engine view takes 17 milliseconds. That results in improvement of 12 milliseconds over 29 milliseconds, which is improvement of 58%.

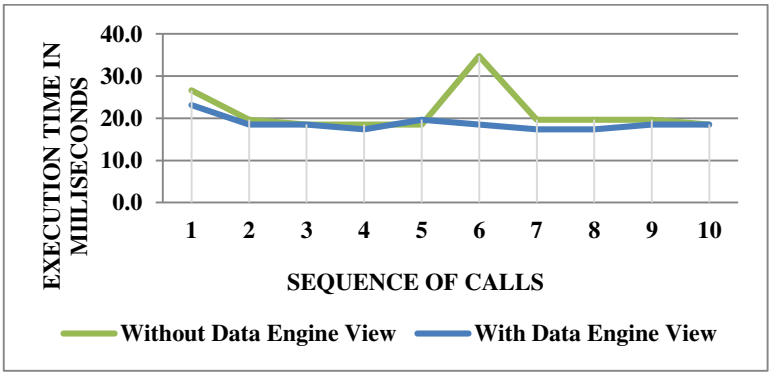
SERVICE CALL PERFORMANCE

From the collected data over larger set of calls, we compare performance of calling individual services in a non-data-engine view with that of calling services packaged as a data-engine service with three different sized of datasets.

Small Data Set - For 10 Records

Graph-4

The collected data are shown below (milliseconds)



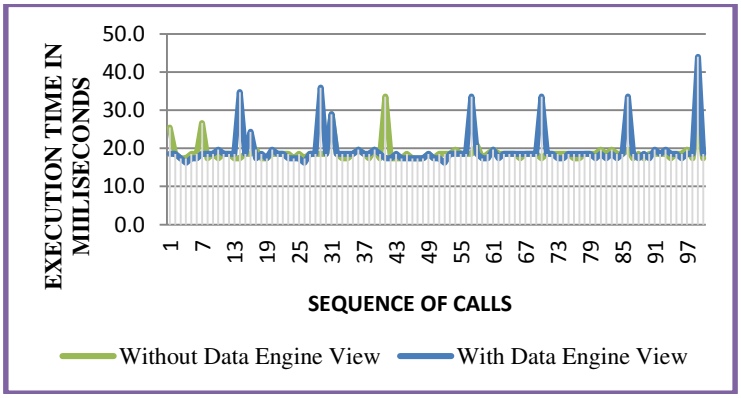
Sources: Authors Compilation

- The calls without data-engine take more time than the calls with data-engine.
- Once the services become more than the sample represented, the time performance is likely to improve further.

Larger Data Set - For 100 Records

Graph-5

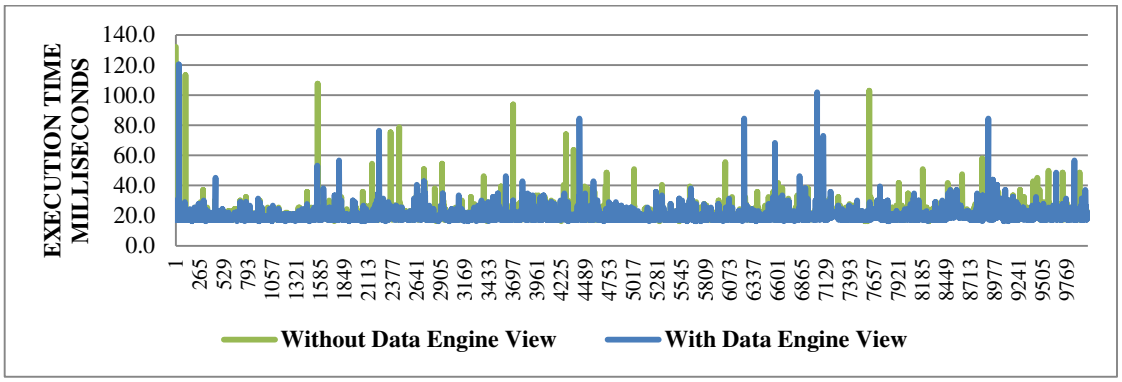
The collected data are shown below (milliseconds)



Sources: Authors Compilation

Larger Data Set - For 10000 Records

Graph-6
The collected data are shown below (milliseconds)



Sources: Authors Compilation

- The calls without data-engine take more time than the calls with data-engine.
- There are instances when calls with data-engine take more time than the calls with without data-engine. It is because of the extra error handling and fault tolerance in data engine where such a case emerges.
- Once the services become more than the sample represented, the time performance is likely to improve further.

CONCLUSIONS

Given analysis of results, clearly assert that a data-engine view gives better performance over fragmented service call structure. For a very large implementation like e-Governance where millions of transactions / records per hour would be created / updated such performance enhancements would be one of core success factors. Data standardization must precede any such architectural exercise.

REFERENCES

1. Papazoglou, P. M. ***Web services: Principles and technology***. ISBN- 9780321155559, Prentice Hall.
2. Murry, J. (2002, October). Designing Manageable Applications. ***Web Developer's Journal***. Retrieved from http://www.webdevelopersjournal.com/articles/design_man_app/
3. Basha (et. al.). (2001). ***A: Professional Java Web Services.***
4. Jeelani, Basha (et. al.). (2001). ***Professional Java Web Services.*** Wrox Press, LIMITED
5. ***JMX technology home page.*** Retrieved from http://java.sun.com/javase/technologies/core/mntrmgmt/javamanagement/
6. Gudgin. (2003, June). ***SOAP 1.2 Part 1: Messaging Framework.*** Retrieved from http://www.w3.org/TR/2003/RECsoap12-part1-
7. Andrews, T. (2003, May). ***Business Process Execution Language for Web Services***. Retrieved from http://www.ibm.com/developerworks/library/ws-bpel
8. ***WS-Eventing specification***. Retrieved from http://www.w3.org/Submission/WS-Eventing/
9. ***SNMP specification***. Retrieved from http://www.sei.cmu.edu/str/indexes/references/SNMPv2_Specs.html
10. Mehta, T. (2003, May). ***Adaptive Web Services Management Solutions***. Retrieved from http://www.enterprisenetworksandservers.com/monthly/toc.php?35
11. Armstrong (et. al.). (2002, March). ***The Java Web Services Tutorial***. Addison-Wesley.
12. Chappell, & Jewell. (2002, March). ***Java Web Services***, O'Reilly.
13. Endrei, M. (2004, April). ***Patterns: Service-Oriented Architecture and Web Services.*** IBM Redbooks SG24-6303-00.

*****

IMAGE CLUSTERING BASED ON COLOR, TEXTURE AND SHAPE

K. Sujatha[5]

ABSTRACT

Image mining is an extension to data mining which uses image processing techniques to extract different patterns related to clustering, classification, association and etc. It is an interdisciplinary field that integrates not only image processing and data mining but also other techniques such as computer vision, machine learning and artificial intelligence. The process of grouping a set of images into classes of similar images without prior knowledge is called image clustering, an important activity in image mining.

This paper explains the process of image clustering by extracting the color, texture and shape features. For color feature extraction RGB, HSV, YCbCr methods, for texture feature extraction GLCM method and for shape feature extraction MPEG-7 Contour Shape descriptor (CSD) are used in this paper.

KEYWORDS

Image Mining, Image Clustering, Data Mining, Colour, Texture, Shape, RGB, HSV, YCbCr, GLCM, Contour Shape Descriptor etc.

INTRODUCTION

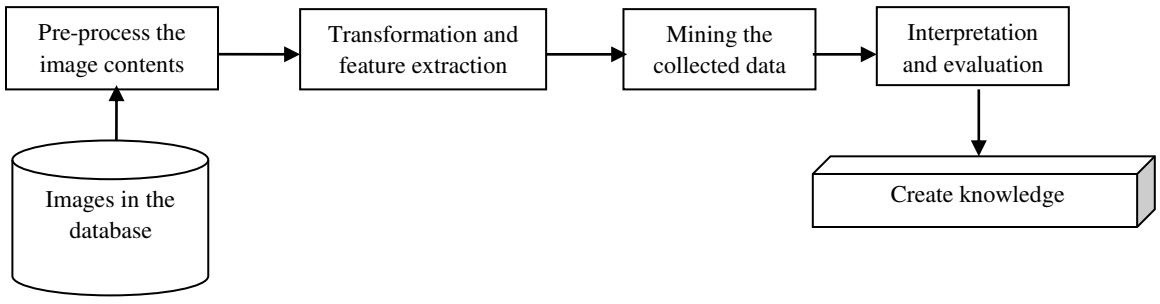
Data mining can be viewed because of the natural evolution of information technology. An evolutionary path has been witnessed in the database industry in the development of the following functionalities: Data collection and database creation, data management (including data storage and retrieval, and database transaction processing), and data analysis and understanding (involving data warehousing and data mining). Data Mining refers to extracting or mining knowledge from large amounts of data [16]. With the emergence of data mining, a large amount of possibilities has been developed to provide some standard solutions to our day-to-day life. In today's world, we can see data mining tasks being performed almost everywhere where the concept of data lies. Therefore, first thing, that we can consider fulfilling our goal is data mining. With this we enter into the promising field of image processing and classification [17].

In this present scenario, image plays vital role in every aspect of business such as business images, satellite images, medical images and so on. If we analysis these data, which can reveal useful information to the human users [11]. Hence, the image mining is rapidly gaining more attention among the researchers in the field of data mining, information retrieval and multimedia databases. Researches can extract semantically meaningful information from image data are increasingly in demand [1].

Image Mining Process

Image mining normally deals with the extraction of implicit knowledge, image data relationship, or other patters not explicitly stored from the low-level computer vision and image processing techniques. (i.e.) the focus of image mining is the in the extraction of patterns from a large collection of images, the focus of computer vision and image processing techniques is in understanding or extracting specific features from a single image.

Figure-1: Image Mining Process [11]



Sources: Authors Compilation

Figure 1 shows the image mining process. The images from an image database are first preprocessed to improve their quality. These images then undergo various transformations and feature extraction to generate the important features from the images.

[5] Associate Professor, CSE Department, Malineni Lakshmaiah Engineering College, Andhra Pradesh, India, sujatha.cb.srinivas@gmail.com

With the generated features, mining can be carried out using data mining techniques to discover significant patterns. The resulting patterns are evaluated and interpreted to obtain the final knowledge, which can be applied to applications [11].

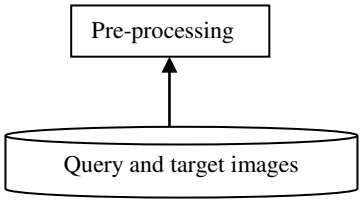
Clustering

Partitioning a set of objects into homogeneous clusters [23] is a fundamental operation in data mining. The operation is needed in a number of data mining tasks, such as unsupervised classification [4] and data summation, as well as segmentation of large heterogeneous data sets into smaller homogeneous subsets that can be easily managed, separately modeled and analyzed [9]. Clustering is a popular approach used to implement this operation. Clustering is the process of grouping the data into classes or clusters so that objects within a cluster have high similarity in comparison to one another, but are very dissimilar to objects in other clusters according to some criteria [5]. Statistical clustering methods [2, 10] use similarity measures to partition objects whereas conceptual clustering methods cluster objects according to the concepts objects carry [6, 15]. The clustering of images is a conceptual clustering, which uses the features of images for clustering. Clustering is the data mining technique, which is often considered as been similar to classification [14]. The similarity between clustering and classification is that both the techniques bring together similar data together. Nevertheless, there is a big difference between them. Clustering can be understood as bringing similar objects into a group called cluster [12]. However, the groups are not predefined. In case of classification, the groups are predefined. Which means in case of Classification what should be part of the group is pre-decided and in case of clustering the grouping is accomplished by finding similarities between data according to characteristics found in the actual data.

Image Retrieval System

For clustering the images, the image must be retrieved and it must be preprocessed.

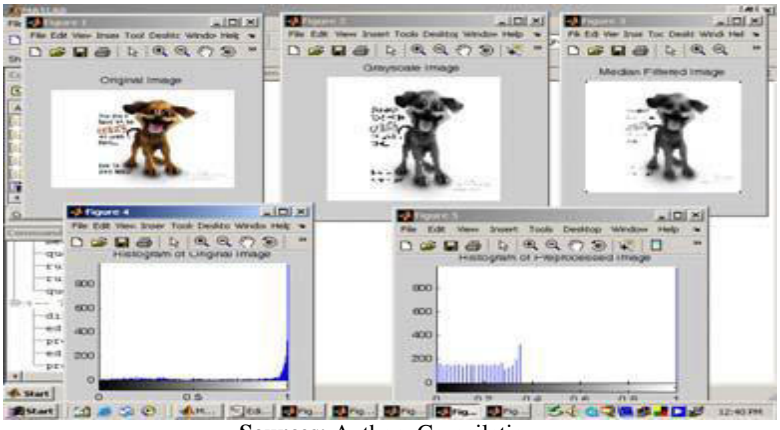
Figure-2: Block Diagram of Image Retrieval System



Sources: Authors Compilation

Pre-processing is the name used for operations on images at the lowest level of abstraction. The aim of the pre-processing is an improvement of the image that suppresses unwilling distortions or enhances some image features, which is important for future processing of the images. This step focuses on image feature processing. Filtering is a technique for modifying or enhancing an image. The image is filtered to emphasize certain features or remove other features. The noise in the images is filtered using linear and non-linear filtering techniques. Median filtering is used here to reduce the noise [13].

Figure-2: Results for Pre-processing Image



Sources: Authors Compilation

REVIEW OF LITERATURE

The image mining and extraction of different patterns such as clustering, classification etc has attracted a great deal of attention since these patterns are used for analyzing images related to many application areas. For developing this paper, number of research papers are studied which are published in different national and international journals and conferences. A. Kannan, Dr. V. Mohan, Dr. N. Anbazhagan [1] has explained about techniques for image retrieval. Anderberg, [2] gives the description about the Cluster Analysis for Applications. B.S. Manjunath, Philippe Salembier, Thomas Sikora [3] provides the description about

MPEG-7 multimedia content description. Zhexue Huang [5] discusses the different data mining algorithms for clustering. G. Foschi Romberg Tiburon, Deepak Kolippakkam*, Huan Liu and Amit Mandvikar [7] gives different methods for feature extraction for image mining. Lokesh Setia, Alexandra Teynor, Alaa Halawani and Hans Burkhardt [12] explain in detail about the texture feature extraction by GLCM. Neelamma K. Patil, Ravi M. Yadahalli, Jagadeesh Pujari [18] gives the descriptions of color feature extraction by RGB, HSV and YCbCr methods. S. Abbasi [19] explains the shape feature extraction. This review of literature is not only limited to these papers but also extended too many research papers represented in references.

RESEARCH METHODOLOGY

In general, images have the following features – color, texture, shape, edge, shadows, temporal details etc. The features that were most promising were color, texture and shape. The reasons are as follows:

- **Color:** Color is the most vital visual feature for humans. By color representation, we mean the overall color of image content when used as a "global" feature. The non-uniformity of RGB color space is eliminated by HSV and YCbCr color space before extracting the features [18].

- **Texture:** Texture is defined as a neighborhood feature [25] – as a region or a block. The variation of each pixel with respect to its neighboring pixels defines texture. Hence, the textural details of similar regions can be compared with a texture template [7].

- **Shape:** Shape may be defined as the characteristic surface configuration of an object: an outline or contour. It permits an object to be distinguished from its surroundings by its outline. Shape representations can be generally divided into two categories [19]: Boundary-based, and Region-based.

Color Feature Extraction

As in many other studies [22, 20, 21, 26], we used color histograms and color co-occurrence matrices to assess the similarity between two images. If the overall color or color pair distributions of two images are close, they are matched as similar in terms of their colors. Three different color spaces are used to produce color histograms; namely RGB, normalized RGB (nRGB) and HSI color spaces. In the RGB color space, each color is represented as a combination of the three primary color channels (Red, Green and Blue). While commonly used, the RGB color space has an important shortcoming, which is the sensitivity to illumination changes. In fact, different color spaces may be suitable in different applications. For instance, the nRGB and the HSI color spaces are often used in order to obtain robustness against illumination differences. The nRGB color model is a derivation of the RGB model in which each channel value is normalized with the total intensity of all channels.

The normalization process effectively normalizes for different illumination conditions. The colors are represented by three normalized color values (nR, nG, nB), which indicate the red, green and blue color ratio in a specific pixel. The normalization computation for red and green channels is formulated as follows: $nR = R/(R+G+B)$ and $nG = G/(R+G+B)$. In the HSI color model, color is represented using its Hue, Saturation and Intensity values. The important feature of this color space is the separation of the intensity from the chromaticity. In order to obtain a histogram robust to normal variations in plant images, the 24-bit RGB information is quantized into a 9-bit representation (for a total of 512 bins, using 3 bits for each color channel). For the nRGB representation, one of the channels can be deduced from the normalized value of the other two ($nR + nG + nB = 1$); therefore, we compute the nRGB color histogram using only the values of two normalized channels, which affords more bins (for a total of 256 bins, using 4 bit for each of the nR and nG values). In the HSI space, the 360 different hue values are quantized to 10, 30 or 90 bins. The intensity value is intentionally discarded, while the saturation component is unused in the current work, for simplicity. Prior to histogram matching, we smooth the computed histograms by taking weighted averages of the consecutive bin values, to obtain some robustness against quantization problems.

Although color co-occurrence is generally mentioned as a texture analysis method, it primarily indicates the distribution of color pairs. We use a 30×30 co-occurrence matrix computed from the HSI color space, where C[i][j] stores the number of neighboring image pixels having the hue values i and j . We generate the co-occurrence matrix using three different methods: (i) considering only four neighboring pixels (i.e. top, bottom, right and left neighbors); (ii) considering all eight neighboring pixels; and (iii) using 8-neighbors but ignoring the diagonal elements of the co-occurrence matrix. Diagonal elements store the number of neighboring pixels that have the same quantized color and dominate the matching process since they correspond to large uniform color regions in the image. Thus, this last method aims to concentrate on areas where colors change, rather than the uniform areas.

Texture Features Extraction

GLCM creates a matrix with the directions and distances between pixels, and then extracts meaningful statistics from the matrix as texture features. [24] GLCM texture features commonly used are shown in the following GLCM is composed of the probability value, it is defined by $P(i,d,\theta)$ which expresses the probability of the couple of pixels at θ direction and d interval. When θ and d is determined, $P(i, j\ d,\theta)$ is showed by $P_{i,j}$. Distinctly GLCM is a symmetry matrix; its level is determined by the image gray-level. Elements in the matrix are computed by the equation showed as follow:

$$P(i,j|d,\theta) = \frac{P(i,j|d,\theta)}{\sum_i \sum_j P(i,j|d,\theta)}$$

GLCM expresses the texture feature according the correlation of the couple pixels gray-level at different positions. It quantification ally describes the texture feature. In this paper, four features is selected, include energy, contrast, entropy, inverse difference.

Energy E= p(x, y)^2

It is a gray-scale image texture measure of homogeneity changing, reflecting the distribution of image gray-scale uniformity of weight and texture.

Contrast I = ΣΣ(x − y)2 p(x, y)

Contrast is the main diagonal near the moment of inertia, which measure the value of the matrix is distributed and images of local changes in number, reflecting the image clarity and texture of shadow depth. Contrast is large means texture is deeper.

Entropy S= ΣΣ p(x, y) log p(x, y)

Entropy measures image texture randomness, when space co-occurrence matrix for all values is equal, it achieved the minimum value; on the other hand, if the value of co-occurrence matrix is very uneven, its value is greater. Therefore, the maximum entropy implied by the image gray distribution is random [8].

Inverse Difference H= Σ (1/(1+(x-y)^2)*P(x,y)

It measures local changes in image texture number. Its value in large is illustrated that image texture between the different regions of lack of change and partial very evenly. Here p(x, y) is the gray-level value at the coordinate (x, y).

Shape Feature Extraction

For retrieving similar images based on shape, the system uses the MPEG-7 Contour Shape descriptor (CSD) defined by Salembier et.al.'s [3] work on MPEG-7 .This CSD is based on the Curvature Scale-Space representation and it captures characteristic shape features of an object in images based on the object contours. This descriptor exploits the fact that the humans tend to decompose shape contours into concave and convex sections while comparing different shapes. Thus, similarity based on shape is assessed in terms of the concave and convex sections, their length relative to the contour length and their position and order on the contour. CSD emulates well the shape similarity perception of the human visual system and is robust to shape deformities. The contour of the object is captured in terms of the eccentricity and circularity values of the original and filtered contours.

Clustering: In this paper Image clustering is done by dividing the process into three parts, namely

Color feature: All Images that are matching with respect to color in database are separated in other folder.

Texture Feature: Later all filtered images matched with respect to texture part and image comparison is done for remaining images, and result is displayed.

Shape Feature: Lastly, images matched in color and texture is extracted with respect to shape and to fulfill the aspect of clustering.

CONCLUSIONS

This paper explains about the clustering of images based on color, texture and shape. In this paper, I used RGB, HSV and YCbCr methods for color feature extraction, GLCM method for texture feature extraction and MPEG-7 contour shape descriptor for shape feature extraction. This paper explains the theoretical explanation for clustering the images based on color, texture and shape. This paper can be extended in future by implementing the process to a real time application and provide the experimental results for clustering and may extended with different methods for feature extractions.

REFERENCES

1. Kannan, A., Mohan, V., & Anbazhagan, N. (2010). Image clustering and retrieval using image mining techniques. ***IEEE International Conference on Computational Intelligence and Computing Research***, ISBN: 97881 8371 362 7.

2. Anderberg, M. R. (1973). ***Cluster Analysis for Applications.*** Academic Press.

3. Manjunath, S. B., Philippe Salembier, Thomas Sikora, (2002). ***Introduction to MPEG- 7: Multimedia Content Description Language.*** Wiley Publications.

4. Cormack, R. M. (1971). A review of classification. ***J. Roy. Statist. Soc.,*** Series A, 134,321-367.

5. Fisher, D. H. (1987). ***Knowledge acquisition via incremental conceptual clustering***, Machine Learning, 2(2), 139-17

6. Markov, I., & Vassilieva, N. (2008). Image retrieval: Color and Texture combining based on query-image ICISP 2008, LNCS 5099. ***Springer- Verlag Berlin Heidelberg***, 430-438.

7. (1996). ***Data Management Solutions***. IBM Corp: IBM White Paper.

8. Jain, A. K., & Dubes, R. C. (1988). ***Algorithms for Clustering Data.*** Prentice Hall.

9. Ji, Zhang Wynne Hsu Mong Li Lee. ***Image mining: Trends and developments***.

10. Lokesh, Setia, Alexandra, Teynor, Alaa, Halawani, & Hans, Burkhardt. (2006, October 26-27). ***Image classification using clustering-cooccurrence matrices of local relational features. California.*** USA: Santa Barbara.

11. Sasikalal, M., & Kumaravel, N. (2005, December11-13). Comparison of feature selection techniques for detection of malignant tumor in brain images. In ***IEEE Indicon 2005 Conference,*** Chennai, India.

12. Margaret, H. Dunham. (2011). ***Data mining introductory and advanced topics*** (7^{th} ed.). 10-13, 119-159, Pearson.

13. Michalski, R. S., & Stepp, R. E. (1983). Automated construction of classifications: conceptual clustering versus numerical taxonomy. ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, 5(4), 396- 410.

14. Michelene, Kamber, & Jiawei, Han. (2001). D***ata mining: Concepts and techniques.*** Simon Fraster University: Academic press.

15. Bhatnagar, Monika, Singh, Kumar Prashant. (2012, March). An efficient technique of image clustering using limited devices***. International Journal of Electronics Communication and Computer Technology (IJECCT),*** ISSN: 2249-7838 IJECCT, 2(2). Retrieved from www.ijecct.org

16. Neelamma, K. Patil, Yadahalli, M. Ravi, Pujari, Jagadeesh. (2011, November). Comparison between HSV and YCbCr color model color-texture based classification of the food grains. ***International Journal of Computer Applications,*** ISSN: 0975-8887, 34(4).

17. Abbasi, S. (1998). ***Curvature scale space in shape similarity retrieval*** (Doctoral Thesis). England: University of Surrey, Centre for Vision, Speech and Signal Processing, Guildford, GU2 5XH.

18. Shim, S. O., & Choi, T. S. (2003, September). Image Indexing by Modified Color Co-occurrence Matrix. ***ICIP***, 2, pp. III- 493–III-496.

19. Sural, S., Qian, G., & Pramanik, S. (2002). Segmentation and histogram generation using the hsv color space for image retrieval. ***ICIP***, 2, pp. II-589–II-592.

20. Tico, M., Haverinen, T., & Kuosmanen, P. (2000, June). A method of color histogram creation for image retrieval. In ***Proc. Nordic Signal Processing Symp. (NORSIG'2000),*** (pp. 157–160). Kolmarden, Sweden.

21. Fayyad, M. U., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. ***Discovery and Data Mining*** (Eds.). AAAI Press/The MIT Press, 573–592.

22. Xiuqi Li, Shu-Ching Chen, Mei-Ling Shyu, & Borko Furht (2002.) Image retrieval by color, texture, and spatial information. In ***Proceedings of the the 8th International Conference on Distributed Multimedia Systems*** (DMS 2002), (pp. 152-159). USA, CA: San Francisco Bay.

23. Y. Rui, T. Huang & S. Chang. (1999, April). Image retrieval: current techniques, promising directions and open issues. ***Journal of Visual Communication and Image Representation***, 10(4), 39-62.

24. Zarit, B., Super, B. &Quek, F. (1999). Comparison of Five Color Models in Skin Pixel Classification. ***In Proc. Int. Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*** (pp. 58-63).

*****

CONCEPTUAL FRAMEWORK OF E-VILLAGE USING INFORMATION COMMUNICATION TECHNOLOGY

Parveen Sharma[6] Prachi Dhaulakhandi[7]

ABSTRACT

Providing basic services like education, water, electricity, employment, health, power, roads to rural areas in developing countries like India, is major issue for the government. A government always tries to set first mile to provide any basic service like Education, if a person can write his name, we can say he or she is literate and the village has 100% literacy rate. This is true even is case of providing ICT services to rural areas. Paper is focused to think about the barriers that force us to be on first mile only in rural ICT services and suggest last mile approach in ICT infrastructure development in rural areas of India. When ICT infrastructure will be well establish, applications of ICT will help the rural population in all basic services. This paper will also suggest a few areas in which ICT can help a lot for rural population of India.

KEYWORDS

ICT, e-Village , e-Governance , WiMax, Gyankendra etc.

THE MEANING OF BEING CONNECTED FOR A VILLAGE

For many years, people working to enhance information & telecommunication infrastructure and applications have referred to rural communities as being at the "last mile of connectivity". However, they are still on "first mile of connectivity". For a rural person, getting connected is a means for sharing the wide range of options available to urbanites, a means for making better and more informed decisions, a means for staying in contact with friends and families who migrate to urban areas for work and education, a means for linking their businesses to the trade, transportation and commerce systems of urban areas, and a means for accessing the services (health, education, information, etc.) that enable urban people to improve their lives.

THE DREAM

To convert a Village into E-Village using ICT, first we have to think, how to motivate the villagers to use the services of ICT. If most of them are living in BPL (Below Poverty Line), first approach should be the economic development providing employment or earning modes to villagers. An awareness of good use of IT is necessary and training to use IT and IT services are required.

POTENTIAL OF ICT FOR RURAL DEVELOPMENT

Rural people constitute the greater part of the population of developing countries and often lack access to basic needs such as water, food, education, health care, sanitation and security. Knowledge and information are basic ingredients for facilitating rural development and bringing about social and economic change. According to Albert Waterson, as quoted by Cohen (1987:23), the purpose of rural development is "to improve the standard of living of the rural population is multi-dimensional including agriculture, industry, and social facilities". Rural communities require information inter alia on supply of inputs, new technologies, early warning systems (drought, pests, and diseases), credit, market prices and their competitors.

Traditional media and ICT have played a major role in diffusing information to rural communities, and have much more potential. There is need to connect rural communities, research and extension networks and provides access to the much needed knowledge, technology and services. Although ICT or the Internet is not a panacea to solve rural development problems, it can open new communication channels that bring new knowledge and information resources to rural communities. Traditional communication channels have been used successfully but these have been monologic or one-way and have not allowed for much interaction with users. Radio for example has been very effective for disseminating information to all types of audiences, but broadcasting times are sometimes not appropriate for most people. However, radio could be linked to the Internet, and a few initiatives have been started on this concept, such as the project Internet Radio in Sri Lanka.

Some examples of areas where ICT could play a catalytic role in economic development for a village in rural area include:

Market outlook – Farmers / Villagers could promote their products and handle simple transactions such as orders using ICT using PDA (A low cost, battery powered small PC, having very simple user interface in local language), linked using BPL (Broadband on Power Lines) or radio link over the web. It is cheaper and faster to trade online than on paper based medium, telephone or fax. e-commerce / m-Commerce could therefore, enable entrepreneurs to access global market information and open up new regional and global markets that fetch better prices and increase farmers' earnings.

Empowering Rural Communities

ICT can empower rural communities and give them "a voice" that permits them to contribute to the development process. With new ICT, rural communities can acquire the capacity to improve their living conditions and become motivated through training

[6] Associate Professor, Department of Computer Science, Manav Bharti University, Himachal Pradesh, India, parveenritika@gmail.com
[7]Research Scholar, CMJ University, Meghalaya, India, prachimanoj8@gmail.com

and dialogue with others to a level where they make decisions for their own development. Giving rural people a voice means giving them a seat at the table to express their views and opinions and become part of the decision making process. The approach should be participatory and could lead to improved policy formation and execution. New ICT have the potential to penetrate under-serviced areas and enhance education through distance learning, facilitate development of relevant local content and faster delivery of information on technical assistance and basic human needs such as food, agriculture, health and water. Farmers can also interact with other farmers, their families, neighbors, suppliers, customers and intermediaries and this is a way of educating rural communities. The Internet can also enable the remotest village to access regular and reliable information from a global library (the web). Different media combinations may be best in different cases through radio, television, video cassettes, audio cassettes, video conferencing, and printouts, CD-ROM or the Internet.

Creating Employment

Through the establishment of rural information centers like Cyberdhabas, Suchanakutir, Kisanghar etc., ICT will create employment opportunities in rural areas by engaging Cyberdhabas, Suchanakutir, Kisanghar, telecentre managers, subject matter specialists, information managers, translators and information technology technicians. Such centers help bridge the gap between urban and rural communities and reduce the rural-urban migration problem. The centers will also provide training and those trained may become small-scale entrepreneurs. That will help in socio-economic development of a village.

ENHANCEMENT IN BASIC SERVICES LIKE EDUCATION, HEALTH, AGRICULTURE, PARTICIPATION

We should mobilize the power of the new media, like the Internet, as well as traditional media such as television, radio and the vernacular newspapers. The combination of the Broadband Internet and community (FM) radio can be particularly powerful to provide better education / awareness timely across to those who need them. We suggest opening **Gyankendra** in villages using various ICT components for rural knowledge centers. The term **Gyankendra** is chosen because at the village level there is need for value addition to generic information by converting it into locale-specific knowledge. With training and technical help in local language, local children, women and men will be able to get basic education that will help us to spread awareness of IT use and applications in rural areas.

LOCAL KNOWLEDGE FOR LOCAL PEOPLE

The relevance of ICT to rural communities will depend upon its ability to provide need-based and local-specific information on health, nutrition, education, natural resource management and livelihood in their own language. Issues relating to weather, water, energy and agriculture, management and marketing of rural products) are important to rural communities.

A ONE-MILLION FREE WORKFORCE

In each of these villages, at least one woman and one man will be selected as Fellows / Motivator of ICT project. We can increase this number up to one million fellows. They will serve as the torchbearers of the knowledge revolution in rural India. It will be useful if the government can proactively promote public-private and institutional partnerships, facilitating the **Gyankendra** centre to cover the villages not reached. The active participation of the elected members of local bodies is crucial to the success of this project. An integrated use of the Internet and community radio will be an effective means of voicing the voiceless.

SUGGESTED FRAMEWORK
Connectivity (Less cost options)

To get connected on High speed and in long range with in less cost the best option is use of radio frequency based broadband (WiMax) , again it can be used easily with wireless devices like Mobile (that are cheap also) ,so we dream a STD /PCO / ITkiosk /VoIP /Mobile ATM /Mobile Health Van with small Information center for a village. Long range and high-speed are two of the key characteristics for connectivity, a radio-waves-based broadband solution that appears to be even better than old technology (Modem, DSL) can be better alternative broadband system approach for remote and rural areas say our E-Village.

- Low cost PC and Other Hardware in Local Language Interface.
- Information Technology awareness in rural areas with the help of students of nearby technology institute as a part of their syllabus.
- Students of B.Tech, MCA, and MBA in thousands of Institutes can be used as free available work force to develop rural market, cheap software and hardware for rural citizens.

REFERENCES

1. Varma, Vasudeva. (2010, January). ***Cloud computing for e-governance***. Hydrabad: IIIT.

2. Kalsi, S. N., Kiran, Ravi & Vaidya, C. S. (2009, January). Effective e-Governance for Good Governance in India. ***International Review of Business Research Papers,*** 5(1), 212-229.

3. Poulose, V. A. (2010, January). E-Governance and Infrastructure: Looking Ahead. ***RITES Journal.***

4. Monga, A. (2008). E-government in India: Opportunities and challenges. ***JOAAG***, 3(2).

*****

A STUDY ON THE AWARENESS, USAGE AND PERCEPTION OF MBA FACULTIES OF ANAND CITY TOWARDS E-BANKING

Snehal J. Bhatt[8] Krishna Gor[9]

ABSTRACT

The fast advancing global information infrastructure (including information technology and computer networks such as the Internet and telecommunications systems) enable the development of electronic commerce at a global level. e-Banking is spreading all over the world, which has changed the attitude of customers towards banking. e-banking is a matter of compulsion for modern day bankers all over the world. However, sometimes e-banking becomes unsafe particularly when it is not backed by improved technology. In India, E-banking is slowing showing its footholds.

This paper aims to determine the usage patterns of E-Banking and perceptions of faculties of MBA course towards E-banking in the city of Anand. For this purpose, 50 faculties of MBA colleges in the Anand city are selected. The questionnaire is sent in the electronic format. The sample collected is analyzed using simple statistics like descriptive statistics analysis, Chi Square, factor analysis etc., the hypothesis were also tested as it deem fit. The study finds that the e-banking facilities used by male and female faculties are not different. The study also finds that the female faculties are more aware about e-banking compared to male faculties.

KEYWORDS

Electronic Commerce, e-Banking, Internet, Perceptions, e-Banking, Anand City etc.

INTRODUCTION

Electronic Banking (E-banking) is an umbrella term for the process by which a customer may perform banking transactions electronically without visiting a brick and mortar institution (FINCEN, 2005). e-banking refers to the effective deployment of IT by banks. It enables the dramatic lowering of transaction costs, and the creation of new types of banking opportunities that address the barriers of time and distance. Customers are no longer wants to wait in queue for paying bills, withdrawing cash etc.

E banking created the opportunities for bank to serve the customer in differentiated way to crate competitive advantage. Further, due to improvement in information technology, customers of banks are now demanding better services from banks at lower rate. E-banking has become now "need" for customer compared to "luxury" as in past. It is affordable to a middle-income person also.

ICICI is the first bank to introduce E banking in India in 1996. Soon various banks like Citi Bank, HDFC Bank and many more followed the route to internet banking. Initially, nationalized banks are reluctant to adopt internet banking; but slowly they also have accepted the internet banking practices and now it is the basic feature of any bank. State Bank of India (SBI) was the first nationalized bank to adopt internet banking in 2001.

India has added 69 million Internet users during 2008-2011 and now has 121 million Internet users with a population penetration rate of 10%, recording a 38% YoY growth. As many as 7% of account holders in the country are using the Internet for banking transactions, while branch banking has fallen by a full 15 percentage points, according to a report by global management consultancy McKinsey & Company.

The following table1 shows the total population using internet and total E banking users in India in 2011. The same data are projected for 2015.

Table-1

Parameter	2011	2015
Indian Population	1.14 Bn	1.27 Bn
Internet Penetration	80 Mn	376 Mn
Internet Penetration %	7.00%	29.50%
Online Shopping	8-10 Mn	38 Mn
Value of ecommerce	$ 6.3 Bn	$ 24 Bn
E-Tailing	2-3 Mn	8 mn
Value of e-tailing	$1.1 Bn	$ 12 Bn
E-tailing penetration of Retail	0.20%	1.41%

Sources: Report by Avendus, 2011

[8]Assistant Professor, Anand Institute of Management, Gujarat, India, snehal.bhatt50@gmail.com
[9]Assistant Professor, Parul Institute of Management & Research, Gujarat, India, gor_krishna@yahoo.com

Table-2: Shows Increase in E-banking w.r.t. ECS, NEFT, Credit cards and Debit cards

GROWING POPULARITY								
Year	ECS (Debit)		NEFT		Credit Cards		Debit Cards	
	Amount	Growth	Amount	Growth	Amount	Growth	Amount	Growth
2007-08	48,937	92	1,40,326	81	57,985	40	12,521	53
2008-09	66,976	37	2,51,956	80	65,356	13	18,547	48
2009-10	69,524	4	4,09,507	63	61,824	-5	26,418	42
2010-11	73,646	6	9,39,149	129	75,516	22	38,691	46

Source: RBI Monthly Bulletin Amount in Rs crore; Growth in %

REVIEW OF LITERATURE

Mahtab Alam and Ankita Soni (2012) studied the customer satisfaction of internet banking with special reference to customers of Vadodara city. 250 internet users were surveyed and the major finding of the study was that there is significant level of difference in the satisfaction level of internet banking users.

Bhagvati R. Pipalyia (2012) studied the customer awareness on internet banking in Gujarat. The finding revealed that there is no significant difference between awareness and usage of E banking.

R. K. Uppal (2011) explored the extent of internet banking in Indian banking industry. The study concludes that the private sector banks are providing internet banking more compared to national counterparts. They also earn high profitability.

A. J. Joshua and Moli P. Koshy (2011) studied the usage patterns of electronic banking services by urban educated customers. The major finding of the study revealed that ATM's are widely accepted but still other electronic banking means are not deployed at the best.

Asli Yuksel Mermod (2011) examines the internet progression in the emerging market like Turkey. The research finds that E banking usage has increased with the increasing education. The major factors considered while choosing internet banking are simplicity of usage and security.

Malhotra Pooja and Singh, B. (2010) made a study on present status of internet banking in India. The result revealed that private and foreign banks provide better service related to E banking compared to public banks.

RESEARCH GAP

The review of literature reveals that most of the studies done on internet banking try to find out the factors influencing the internet banking, perceptions and awareness about E banking, satisfaction level with respect to E banking among the customers in India and other countries. The present study will try to add to the existing literature and find out the usage and perception of faculties of MBA courses in the city of Anand with respect to E banking.

OBJECTIVES OF STUDY

1. To study the awareness of E banking among the faculties of MBA course in the city of Anand.
2. To identify the factors considered by faculties of MBA course while selecting E banking.
3. To find out the major E banking facilities used by faculties of MBA course in the city of Anand.

RESEARCH METHODOLOGY

Hypothesis of Study

H_1: There is significant difference in awareness among the male and female faculties of MBA course with respect to E banking in the city of Anand.

Data Collection

Data were collected from 30 faculties of MBA course who have bank accounts in the city of Anand during October – November 2012. Purposive sampling is used for selection of samples. The survey instruments used is structured questionnaire. There are two parts of the questionnaire: first part is related to research questions and second part is related to personal information. Analysis is done by using percentage method and software's like SPSS.

DATA ANALYSIS AND INTERPRETATIONS

Table-3: Demographic Profile

Parameter	Range	Male		Female	
		Frequency	Percentage	Frequency	Percentage
Gender	Gender	25	50%	25	50%
Marital Status	Married	20	71%	19	70%
	Unmarried	5	29%	76	30%
Age	20 to 30 Years	9	36%	8	32%
	31 to 40 Years	6	24%	7	28%
	41 to 50 Years	5	20%	5	20%
	>50 Years	5	20%	5	20%
Teaching Experience	1 to 3 Years	12	48%	12	48%
	3 to 5 Years	7	28%	8	32%
	5 to 10 Years	3	12%	4	16%
	> 10 Years	3	12%	1	4%
Education	P.G	20	80%	22	88%
	Doctorate	5	20%	3	12%
Income (p.m.)	10,000 to 20,000	12	48%	15	60%
	20,000 to 30,000	5	20%	7	28%
	30,000 to 40,000	3	12%	1	4%
	> 40,000	5	20%	2	8%

Sources: Primary Data Collection

Hypothesis:

H_0: There is no significant difference in awareness among the male and female faculties of MBA course with respect to E banking in the city of Anand.

Table-4: **Mean Ranks of Awareness of Male and Female Faculties**

	Gender	N	Mean Rank	Sum of Ranks
Rate the awareness of E banking	Male	25	23.00	575.00
	Female	25	28.00	700.00
	Total	50		

Sources: Primary Data Collection

From the above table, it can be interpreted that the mean ranks of female faculties is more than mean rank of male faculties and hence female faculties are more aware about E banking facilities.

Table-5: Test Statistics for Mann Whitney Test

	Rate the Awareness of E banking
Mann-Whitney U	250.000
Wilcoxon W	575.000
Z	-1.271
Asymp. Sig. (2-tailed)	.204

Sources: Primary Data Collection
a. Grouping Variable: Gender

At $\alpha = 0.05$ level of significance, there is enough evidence to conclude that there is no significant difference in awareness among the male and female faculties of MBA course with respect to E banking in the city of Anand. So, Null hypothesis is not accepted.

FINDINGS

- Majority of faculties (86%) are aware about E banking facilities provided by their banks.
- 65% of the faculties have their account in private banks and rest 35% has their account in public banks.
- Majority of faculties (76%) are using E banking facilities provided by their bank.
- 85% faculties use E banking for money/fund transfer.
- 70% of faculties are using E banking from last 1 year only.
- 80% faculties believe that prior knowledge of computer is required for using E banking transactions.

- 86% faculties agree that E banking is easy to use as it can be operated from home also.
- Majority of faculties (78%) agrees that using E banking facilities is secure and there transaction details are kept secret by using E banking.
- 83% faculties agree that using E banking will give them respect from the peers and their peers would consider them as techno savvy person.
- 95% faculties' agrees that their bank provides the recent updates to them timely.

CONCLUSIONS

It can be concluded from the study that nearly all the faculties are using E banking. Females are more aware than then male faculties regarding E banking. There is no specific pattern analyzed regarding the usage of the facilities provided by the bank among males and females. Faculties' believes that E banking is the safe mode of operation and the financial transaction details are kept secret by using E banking. It is also observed that faculties who are using of E banking believe that it will help to get respect from the peers and will project them as techno savvy person.

REFERENCES

1. Shanker, Ravi. (2003). ***Services Marketing – the Indian perspective***. Excel Books.
2. (2010, February). Banking sector in India. ***Yojana***.
3. Retrieved from http://www.financialexpress.com/news/online-banking-users-in-india-now-close-to-a-million-says-idc/52305/
4. Retrieved from http://www.medianama.com/2012/06/223-india-has-121-million-internet-users-39-million-3g-subscribers-report/
5. Retrieved from http://www.bankingonly.com/detail-news.php?news_id=665
6. Retrieved from http://www.123eng.com/forum/f12/recent-trends-internet-banking-8574/
7. Retrieved from http://www.arraydev.com/commerce/jibc/2011-04/JoshuaAmbat.pdf
8. Retrieved from http://www.primejournal.org/BAM/pdf/2011/mar/Uppal%20pdf%202.pdf
9. Retrieved from http://philica.com/display_article.php?article_id=320
10. Retrieved from http://www.arraydev.com/commerce/jibc/2011-04/YukselMermodAsl.pdf

*****

A MODEL FOR IDENTIFYING GUILTY AGENTS IN DATA TRANSMISSION USING A FAKE OBJECT INJECTION

Sumit Agarwal[10]

ABSTRACT

Aim of this research paper is to detect when the PC distributor is sensitive their agents have leaked data, and if possible, to identify the agent that leaked the data. PC distributor has given sensitive data to a set of purportedly trusted agents. Some of the data is leaked and found in a third party place (e.g., on the web or somebody's laptop). The distributor must assess the likelihood that the leaked data came from one or more agents, as opposed to having been independently gathered by other means. Data allocation strategies (across the agents) are proposed that improve the probability of identifying leakages. These methods do not rely on alterations of the released data (e.g., watermarks). In some cases, the distributor can also inject "realistic but fake" data records to further improve our chances of detecting leakage and identifying the guilty party. A model for assessing the "guilt" of agents using C# dot net technologies with MS-SQL server as backend is proposed to develop. Algorithms for distributing objects to agents, in a way that improves our chances of identifying a leaker is aloes presented. Finally, the option of adding "fake" objects to the distributed set is also considered. Such objects do not correspond to real entities.

KEYWORDS

Data Leakage, Data Privacy, Fake Record, Unobstrstuive, Agents, Techniques etc.

INTRODUCTION

While doing business, sometimes the sensitive data must be handed over to supposedly trusted third parties. For example, a hospital may give patient records to researchers who will devise new treatments. Similarly, a company may have partnerships with other companies that require sharing customer data. Another enterprise may outsource its data processing, so data must be given to various other companies. We call the owner of the data the distributor and the supposedly trusted third parties the agents.

Our goal is to detect when the distributor is sensitive agents have leaked data, and if possible, to identify the agent that leaked the data. We consider applications where the original sensitive data cannot be perturbed. Perturbation is a very useful technique where the data is modified and made "less sensitive" before being handed to agents. For example, one can add random noise to certain attributes, or one can replace exact values by ranges. However, in some cases it is important not to alter the original distributor's data. For example, if an outsourcer is doing our payroll, he must have the exact salary and customer bank account numbers. If medical researchers will be treating patients (as opposed to simply computing statistics), they may need accurate data for the patients.

Traditionally, leakage detection is handled by watermarking, e.g., a unique code is embedded in each distributed copy. If that copy is later discovered in the hands of an unauthorized party, the leaker can be identified. Watermarks can be very useful in some cases, but again, involve some modification of the original data. Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious.

In this work, unobtrusive techniques for detecting leakage of a set of objects or records have been studied. For example, after giving a set of objects to agents, the distributor discovers some of those same objects in an unauthorized place. At this point, the distributor can assess the likelihood that the leaked data came from one or more agents, as opposed to having been independently gathered by other means. Using an analogy with cookies stolen from a cookie jar, if we catch Freddie with a single cookie, he can argue that a friend gave him the cookie. However, if we catch Freddie with five cookies, it will be much harder for him to argue that his hands were not in the cookie jar. If the distributor sees "enough evidence" that an agent leaked data, he may stop doing business with him, or may initiate legal proceedings.

In this paper, we develop a model for assessing the "guilt" of agents. We also present algorithms for distributing objects to agents, in a way that improves our chances of identifying a leaker. Finally, we also consider the option of adding "fake" objects to the distributed set. Such objects do not correspond to real entities but appear realistic to the agents. In a sense, the fake objects acts as a type of watermark for the entire set, without modifying any individual members. If it turns out an agent was given one or more fake objects that were leaked, then the distributor can be more confident that agent was guilty.

EXISTING SYSTEM

Perturbation

Application where the original sensitive data cannot be perturbed has been considered. Perturbation is a very useful technique where the data is modified and made "less sensitive" before being handed to agents. For example, one can add random noise to certain attributes, or one can replace exact values by ranges. However, in some cases it is important

[10] Assistant Professor, Department of Computer Science, Vishveshwarya Group of Institutions, Uttar Pradesh, India, sumitcomputer@gmail.com

not to alter the original distributor's data. For example, if an outsourcer is doing our payroll, he must have the exact salary and customer bank account numbers. If medical researchers will be treating patients (as opposed to simply computing statistics), they may need accurate data for the patients.

Watermarking
Traditionally, leakage detection is handled by watermarking, e.g., a unique code is embedded in each distributed copy. If that copy is later discovered in the hands of an unauthorized party, the leaker can be identified. Watermarks can be very useful in some cases, but again, involve some modification of the original data. Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious.

Disadvantages

Consider applications where the original sensitive data cannot be perturbed. Perturbation is a very useful technique where the data is modified and made "less sensitive" before being handed to agents. However, in some cases it is important not to alter the original distributor's data. Traditionally, leakage detection is handled by watermarking, e.g., a unique code is embedded in each distributed copy. If that copy is later discovered in the hands of an unauthorized party, the leaker can be identified. Watermarks can be very useful in some cases, but again, involve some modification of the original data. Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious.

PROPOSED SYSTEM

Unobtrusive techniques for detecting leakage of a set of objects or records have been studied. After giving a set of objects to agents, the distributor discovers some of those same objects in an unauthorized place. (For example, the data may be found on a web site, or may be obtained through a legal discovery process.) At this point, the distributor can assess the likelihood that the leaked data came from one or more agents, as opposed to having been independently gathered by other means.

Using an analogy with cookies stolen from a cookie jar, if Freddie with a single cookie has been cached, he can argue that a friend gave him the cookie. However, if Freddie with five cookies has been cached, it will be much harder for him to argue that his hands were not in the cookie jar. If the distributor sees "enough evidence" that an agent leaked data, he may stop doing business with him, or may initiate legal proceedings.

A model for assessing the "guilt" of agents has been developed. An algorithm for distributing objects to agents, in a way that improves our chances of identifying a leaker has been proposed. The option of adding "fake" objects to the distributed set also been considered. Such objects do not correspond to real entities but appear realistic to the agents. In a sense, the fake objects acts as a type of watermark for the entire set, without modifying any individual members. If it turns out an agent was given one or more fake objects that were leaked, then the distributor can be more confident that agent was guilty.

Advantages

After giving a set of objects to agents, the distributor discovers some of those same objects in an unauthorized place. At this point, the distributor can assess the likelihood that the leaked data came from one or more agents, as opposed to having been independently gathered by other means. If the distributor sees "enough evidence" that an agent leaked data, he may stop doing business with him, or may initiate legal proceedings to develop a model for assessing the "guilt" of agents. We also present algorithms for distributing objects to agents, in a way that improves our chances of identifying a leaker. Consider the option of adding "fake" objects to the distributed set. Such objects do not correspond to real entities but appear. If it turns out an agent was given one or more fake objects that were leaked, then the distributor can be more confident that agent was guilty.

IMPLEMENTATION PLAN

In this application, we try to implement a model application to detect the data leakages between distributor and agents. The system is developed in C# dot net. We use Microsoft SQL Server 2000 as database for this application. When the distributor sends a file to agent, it is considered as fake object for that particular agent, in this sequence file name and file path is stored in the database for future reference.

Similarly, when the agent sends a file to the unauthorized agent the sequence is store in the database. Thus, we can find the guilt agent. The probability function is calculated based on the number of guilt agents by the number of file transfers between the agent and unauthorized person.

MODULE DESCRIPTION

1- **Login / Registrations**
This is a module mainly designed to provide the authority to a user in order to access the other modules of the project.

2- **Data Transfer**
This module is mainly designed to transfer data from distributor to agents. The same module can also be used for illegal data transfer from authorized to agents to other agents

3- **Guilt Model Analysis**
This module is designed using the agent – guilt model. Here a count value (also called as fake objects) is incremented for any transfer of data occurrence when agent transfers data. Fake objects are stored in database.

4- **Agent Guilt Model**
This module is mainly designed for determining fake agents. This module uses fake objects (which is stored in database from guilt model module) and determines the guilt agent along with the probability. A graph is used to plot the probability distribution of data, which is leaked by fake agents. To compute this probability, we need an estimate for the probability that values can be "guessed" by the target.

ALGORITHM STEPS

Step1 - Distributor select agent to send data.
The distributor selects two agents and gives requested data R1, R2 to both agents.

Step2 - Distributor creates fake object and allocates it to the agent
The distributor can create one fake object (B = 1) and both agents can receive one fake object (b1 = b2 = 1). If the distributor is able to create more fake objects, he could further improve the objective.

Step3- Check number of agents, who have already received data.
Distributor checks the number of agents, who have already received data.

Step4- Check for remaining agents.
Distributor chooses the remaining agents to send the data. Distributor can increase the number of possible allocations by adding fake object.

Step5- Select fake object again to allocate for remaining agents.
Distributor chooses the random fake object to allocate for the remaining agents.

Step6- Estimate the probability value for guilt agent to compute this probability, we need an estimate for the probability that values can be "guessed" by the target.

WHY USE DATAMINING?

Data mining is the process of extracting patterns from data. Data mining is becoming an increasingly important tool to transform the data into information. It is commonly used in a wide range of profiling practices such as marketing, surveillance fraud detection and scientific discovery. Data mining can be used to uncover patterns in data but is often carried out only on samples of data. The mining process will be ineffective if the samples are not a good representation of the larger body of data. Data mining cannot discover patterns that may be present in the larger body of data if those patterns are not present in the sample being "mined". Inability to find patterns may become a cause for some disputes between customers and service providers. Therefore, data mining is not foolproof but may be useful if sufficiently representative data samples are collected. The discovery of a particular pattern in a particular set of data does not necessarily mean that a pattern is found elsewhere in the larger data from which that sample was drawn. An important part of the process is the verification and validation of patterns on other samples of data

REVIEW OF LITERATURE

The guilt detection approach we present is related to the data provenance problem, tracing the lineage of 'S' objects implies essentially the detection of the guilty agents.

- Suggested solutions are domain specific, such as lineage tracing for data warehouses and assume some prior knowledge on the way a data view is created out of data sources.

- Watermarks were initially used in images, video and audio data whose digital representation includes considerable redundancy. Watermarking is similar in the sense of providing agents with some kind of receiver-identifying information. However, by its very nature, a watermark modifies the item being watermarked. If the object to be watermarked cannot be modified then a watermark cannot be inserted. In such cases, methods that attach watermarks to the distributed data are not applicable.

- Recently, works have also studied marks insertion to relational data.

There are also many other works on mechanisms that allow only authorized users to access sensitive data through access control policies. Such approaches prevent in some sense data leakage by sharing information only with trusted parties. However, these policies are restrictive and may make it impossible to satisfy agents' requests.

ALGORITHM RESULTS TO FIND OUT GUILTY AGENTS

Figure-1

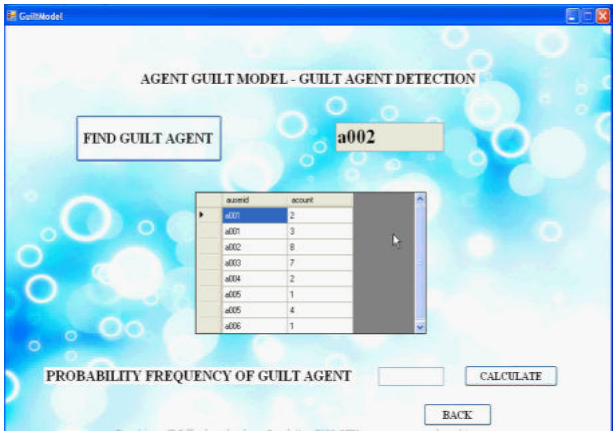
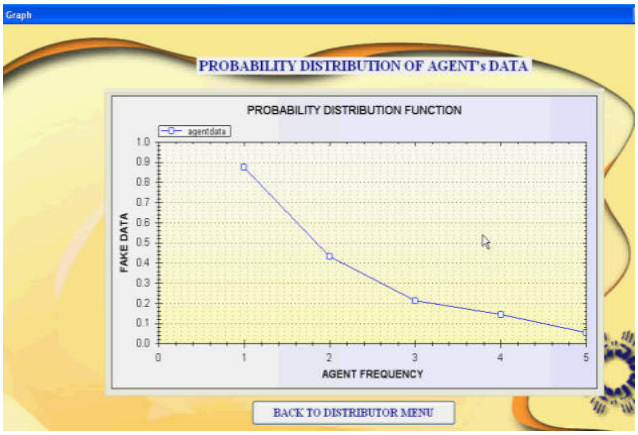


Figure-2



Sources: Authors Compilation & taken by execution of proposed Algorithm in .Net Technology

We implemented algorithms in .Net Frame work using C# .The above figure are the results of execution .Figure 1 shows the list of all guilty agents and Figure 2 shows the probability of agent's data.

CONCLUSIONS

The likelihood that an agent is responsible for a leak is assessed, based on the overlap of his data with the leaked data and the data of other agents, and based on the probability that objects can be "guessed" by other means. The algorithms we have presented implement a variety of data distribution strategies that can improve the distributor's chances of identifying a leaker. We have shown that distributing objects judiciously can make a significant difference in identifying guilty agents, especially in cases where there is large overlap in the data that agents must receive.

REFERENCES

1. Agrawal, R., & Kiernan, J. (2002). Watermarking relational databases. In VLDB '02, ***Proceedings of the 28th international conference on Very Large Data Bases*** (pp. 155–166). VLDB Endowment.

2. Bonatti, P., Vimercati, Di C. D. S., & Samarati, P. (2002). Algebra for composing access control policies. ***ACM Trans. Inf. Syst. Secur.,*** 5(1), 1-35.

3. Buneman, P., Khanna, S. & Tan, C. W. (2001, January 4-6). Why and where: A characterization of data provenance. In J. V. den Bussche and V. Vianu (Ed.), ***Proceedings of Database Theory - ICDT 2001, 8th International Conference***, (Vol. 1973).London, UK.

4. Buneman P. & Tan, C. W. (2007). Provenance in databases. In SIGMOD '07, ***Proceedings of the 2007 ACM SIGMOD international conference on Management of data*** (pp. 1171-1173). ACM, New York, NY: USA.

5. Cui, Y., & Widom, J. (2001). Lineage tracing for general data warehouse transformations. ***The VLDB Journal,*** 471-480.

6. Murthy, N. V. (1981). Counting the integer solutions of a Linear Equation with Unit Coefficients. ***Math Magazine,*** 54(2), 79-81.

7. Nabar, U. S., Marthi, B., Kenthapadi, K., Mishra, N., & Motwani, R. (2006). Towards Robustness in Query Auditing. In (VLDB '06), ***Proc. 32nd Int'l Conf. Very Large Data Bases*** (pp. 151-162). VLDB Endowment.

8. Papadimitriou, P., & Garcia-Molina, H. (2008). ***Data Leakage Detection*** (Technical Report). Stanford Univ.

9. Sion, R., Atallah, M. & Prabhakar, S. (2003). Rights Protection for Relational Data. In ***Proc. ACM SIGMOD*** (pp.98-109).

10. Sweeney, L. (2002). ***Achieving k-anonymity privacy protection using generalization and suppression.*** Retrieved from http://en.scientificcommons.org/43196131

*****

VIRAL MARKETING OF DIGITAL PRODUCTS USING SOCIAL MEDIA

K. Murugan[11] P. Venkatesh[12]

ABSTRACT

Improvements in hardware and software technologies like high-speed internet, cloud computing, smaller and faster chips, have made social networking and mobile devices ubiquitous, which has in turn created a huge opportunity in digital products and services market. Marketers, in trying to use traditional word-of-mouth marketing concepts online for viral marketing for their digital products are realizing that there is the potential of exponential growth that can be achieved very quickly and very cheaply when compared to using more traditional marketing channels. This report attempts to explore how marketers could use viral marketing to market their digital products and realize this potentially exponential growth.

In answering this question, this report draws on results of primary and secondary research, including four interviews conducted in March and April 2011 with professionals from organizations dealing in digital products and social media marketing. Among the topics covered in these interviews were identifying who can use viral marketing, strategic issues surrounding viral marketing, specific characteristics that a products needs to have to be considered for viral marketing, creating and executing a viral marketing campaign, and how to make a viral marketing campaign sustainable.

The research showed that all aspects of an organization need to come together and work in tandem to potentially achieve an exponential growth using a viral marketing campaign - from defining an overall business and marketing strategy, looking at company's capabilities, putting crisis management in place, developing the right product which is social spread friendly, finding the right influencers in the relevant market channels, seeding these influencers, monitoring the campaign, engaging with customers as they provide positive and negative feedback, and all this while building momentum to a point where campaign potentially goes viral.

KEYWORDS

Viral Marketing, Software, Digital Products, Social Media etc.

INTRODUCTION

Digital Products

In 2002, Hui et al (Hui & Chau 2002) classified Digital products into three categories:

- Tools and Utilities that assist user to accomplish specific goals or tasks,
- Content based products whose value lies in the information content,
- Online services that provide access to useful resources like server connections as well as online utilities that assist users in accomplishing specific tasks.

Since then, with the rise and adaptation of Web 2.0 technologies, digital products and services have seen an explosion - both in terms of the numbers of products and services, and the number of consumers using them i.e. mobile platforms (Facebook, Twitter, Blogs, YouTube, LinkedIn etc.), mobile applications (>300,000 apps on Apple AppStore with 10 billion downloads), Cloud computing etc.

Social media and Viral Marketing

Internet has now become a virtual social world where we meet our friends and family, make new friends, engage in conversation about work-life experience – including buying products and services and the resulting experiences – engage in marketplace and listen to people we trust when making buying decisions.

REVIEW OF LITERATURE

Social Media?

Social Media is the commonly known term for the Web2.0 technologies that enable the Internet users to generate and exchange content (Kaplan & Haenlein 2010) through desktop and mobile devices. With the increase in consumption of the mass media and internet, the decline of community activity has been one of the dominant social trends of recent decades across the world's advanced economies (Putnam 2000). American social scientist Robert Putnam wrote about this trend, but also saw the potential of revival of these communities through internet (Putnam 2000). The Cluetrain Manifesto (Levine 2009), the 1999 internet

[11] Assistant Professor, Sri Sairam Institute of Management Studies, Sri Sairam Engineering College, Tamil Nadu, India, muruganindia2002@yahoo.com
[12] Assistant Professor, Sri Sairam Institute of Management Studies, Sri Sairam Engineering College, Tamil Nadu, India, venky.rose@yahoo.co.in

marketing book also made a similar point in that it claimed people were drawn to the internet because of "the promise of voice and thus of authentic self (Levine 2009)." Social media's popularity (Mashable 2011) over the last few years seems to provide confirmation to this as people continue to reach out to connect to other human beings and in the process accept the technological advancements that are thrown at them.

The term Social Media was coined around 2004 with the launch of social networking websites like MySpace (2003) and Facebook (2004); however, the concept of social networking can be traced back to almost 30 years ago when the first email was sent between two computers. Advancements in technologies like Internet, network infrastructure, Web 2.0, mobile devices etc. have since empowered every internet user to create their own content - be it video, blog, opinions etc. - and share between their online social networks. Social Media allows people to stay connected with many more people – friends, families, business associates etc., - even across multiple continents, than was possible in the past.

What does Social Media means for marketers?

Gerzema and Lebar (Gerzema & Edward 2008) argue in their book The Brand Bubble that since consumers trust each other more than they trust marketing information, social media has altered the trust equation for brands by allowing the customers to create and exchange their own contents. This means the customers can freely exchange positive and negative perceptions about a brand in a connected world where these perceptions can spread like a wild fire. For example, Groupon's 2011 Superbowl advertisement was met by huge backlash on Twitter, so much so that the CEO of Groupon had to explain their reasoning behind choosing those specific advertisements on the company blog (Andrew 2011). The technological advancements, increased connectivity online, and offline have also allowed an unprecedented number of new brands to be introduced in recent years globally (Gerzema & Lebar 2009). Gerzema & Lebar 2009, discovered that consumer attitudes towards all sizes and segments of brands were in serious decline. They observed significant drops in key measures of brand value - 'top of mind' awareness, trust, regard and admiration – aka the Brand Equity.

WHAT IS VIRAL MARKETING?

Viral Marketing is the intentional influencing of consumer-to-consumer communication by professional marketing techniques (Kozinets et al. 2010). It is also otherwise known as word- of-mouth marketing (WOMM), buzz marketing and guerrilla marketing. Wilson (Wilson 2005) describes viral marketing as:

> "Any strategy that encourages individuals to pass on a marketing message to others, creating the potential for exponential growth in the message's exposure and influence".

Advantages and Disadvantages of VM

When looking to buy products/services almost 76% people rely on other people's opinions for product recommendations, versus 15% on advertising (Qualman 2009). Yang et al (Yang& Allenby 2003) showed that the geographically defined network of consumers is more useful than the demographic network for explaining consumer behaviour in purchasing Japanese cars. Hill et al (Hill, Provost & Volinsky 2006) found that adding network information, specifically whether a potential customer was already talking to an existing customer, was predictive of the chances of adoption of a new phone service option. For the customers linked to a prior customer, the adoption rate was 3–5 times greater than the baseline. These recommendations come from people who are opinions leaders or someone who has influence within a community. In order to capture the attention of these influencers, marketers have to use targeted marketing campaigns.

VM campaign is a lower-cost option when compared to mass media marketing. It also allows the marketers to target specific customers and has a high and rapid response rate. It appears to be quite a straightforward option for the organizations and marketers to adopt viral marketing strategies; however, the outcome of VM is hard to predict (Watts, Peretti & Frumin 2007). Mass marketing, on the other hand, has a far wider reach and marketers can get predictable returns from such campaigns. Watts et al (Watts, Peretti & Frumin 2007) argue that combining viral marketing with mass marketing, in what they call Big Seed Marketing, would allow marketers to get a more predictable return. Marketers should also carefully consider the fact that viral marketing campaigns become unmanageable once they gain a certain momentum. This can be a major problem for marketers, especially if the campaign does not have the desired outcome.

In a study in 2001, Bowman et al (Bowman & Narayandas 2001) found that self-reported loyal customers were more likely to talk to others about the products when they were dissatisfied, but, interestingly, they were not more likely to talk to others when they were satisfied. To avoid this, marketers need to carefully monitor the early stages of a VM campaign when they may have some control and there is an opportunity to take corrective measures. The low cost of social media marketing and the potential to reach and influence millions of people through personal recommendation seems to be a very lucrative option for marketers. However, if not managed properly, such campaigns also have the potential to get out of hand very quickly and become unmanageable. Marketers need to keep this mind and take a balanced approach.

MEASURING EFFECTIVENESS OF A VM CAMPAIGN

Emerging social network analysis and visualization techniques offer the marketers to delve deeper into consumer minds (Whitney 2010) – to identify connectors, influencers, implementers and other types of members in the group. Whitney (Whitney 2010) describes that social network visualizations can help identify important connecting points such as pre- established relationships,

shared expertise, and who may have information that is not obvious from their current roles. Event Graphs (Hansen, Smith & Shneiderman 2010) can be used to illustrate the structure of connections and communications among people discussing an event.

These Event Graphs can help identify sub-groups within larger conversations, as well as individuals with unique social signatures. Jesse Thomas and Brian Solis (Solis & Thomas 2011), the social media gurus, recently developed an info graphic of the Twitter verse depicting important tools to help marketers more effectively navigate, engage, analyze and measure participation on Twitter.

RESEARCH METHODOLOGY

This section provides an outline of the research methodology used to answer the research questions - the research approach, a description of primary data collection process for the interviews, secondary research, and data analysis techniques used and limitations of the adopted research method.

Research Approach

The research approach influences design and provides an opportunity to consider benefits and limitations of various approaches available to the researcher (Crewell 2003). Two types of approaches are available – deductive and inductive. Deductive approach tests theories, while an inductive approach forms theories (Marcoulides 1998). This report uses inductive research approach as it aims to formulate hypothesis and develop general theory around how organizations could go about marketing digital products, especially viral marketing, using social media. There are two methods available for data analysis – Qualitative and Quantitative. Qualitative research is "a research strategy that usually emphasizes words rather than quantification in the collection and analysis of data" (Bryman & Bell 2007), while quantitative research is based on data analysis to generate reliability.

Qualitative research better reflects, "The quality of the lived experience of individuals, which cannot be reduced to numerical values using statistical analysis" (Hewitt-Taylor 2001). Social media is a dynamic field, which is continuously changing. This means that although quantitative analysis would provide data, designing an appropriate survey to get qualified opinions and understand the deeper issues in this area was quite challenging (Amaratunga et al. 2002). This report uses the qualitative method to explore the research questions as it allows researchers to conduct in-depth explorations of a particular phenomenon (Crewell 2003). This choice is further justified as the research questions focus on opinions, feelings and experiences, thus providing subjective data.

Data Collection Methods

In-depth interviews act as the primary source of the research and syndicate services (Twitter, blogs, Facebook groups etc.) as the secondary sources.

Qualitative analysis allows for a better understanding and interpretation of the experiences of their subjects (Tvede & Ohnemus 2001). This is important for this research as it looks at a wide variety of experiences, understanding of the subject and interpretation of the data gathered. In order to achieve this, in-depth personal interviews with market participants were conducted at prearranged locations. The discussion in the interviews was structured around the core research questions, but no set questionnaire was developed. The interviews themselves were semi-structured in that even though the questions were based on research questions, they were kept open ended and the direction of the discussions was based on the interviewee's experience and area of expertise.

The reasons for selecting semi-structured interviews as the preferred approach are:

- They involve a series of open-ended questions allowing the discussion on research topics.
- They allow the interviewer to encourage the interviewee to consider a question further.
- They provide a high level of response.

Necessary precautions were taken to ensure that there were no faults in recording the interviews. Interviews were recorded on an audio recording device and then carefully transcribed to avoid any such issues.

The questions in the interviews were sequenced in the following conceptual order with related to social media strategy and capability assessment:

- Marketing strategy, especially the importance of an integrated approach
- Design attributes for Product / Service for them to be considered for Viral Marketing
- Designing a viral marketing message for product/services
- Importance of influencers

- Executing, managing and monitoring a viral marketing campaign, and the importance of sentiment analysis
- Exploring whether a viral marketing campaign can be made sustainable

Candidates for the interview were selected based on their experience in social media marketing and their relevance in the managing overall strategy and executing specific parts of SMM strategy.

Due to geographical constraints and professional commitments, one of the interviews - with the Online Community Manager of KIK Interactive Inc - had to be done as email conversations. It would have been better to conduct this interview as a telephone conversation but it was not feasible because of the time differences and other work commitments on their part. All interviews were conducted in March 2011.

SWOT Analysis

Strength:
Starting with an online word of mouth. Hence provides effortless transfer op products to others and takes advantages of others.

Weakness:
Messages can be treated as spam and brand dilution

Opportunity:
It allows sharing, downloading and embedding, and connects with comments.

Threats:
It has associations with unknown persons.

Interview Subjects

Colin Gilchrist: www.socialtailor.com: Colin is a social media strategist, who helps organizations assess their overall marketing strategy and help them integrate social media marketing as part of this strategy.

Andrew Burnett: Urban Niche: Andrew helps companies design and executes social media marketing campaigns. Andrew's company specializes in pushing the marketing message with a view to reach a point where the message could go viral.

Secondary Sources

The rapid changes in the field of social media mean that there is a scarcity of academic literature in this area. Hence, opinion leaders and practitioners in the field of social media marketing focused a lot of research on works. In order to get an appropriate range of secondary resources, wide reading was done including these sources: a) Analysts reports, b) Industry and academic journals, c) Blogs, d) Twitter, e) Facebook groups, and f) White papers.

Witter turned out to be one of the most useful resources for secondary research, as it seems to have become a platform where all the latest ideas are shared in real time, based on the experiences of the companies and the thought leaders involved in this area.

DATA ANALYSIS

The data processing in this report is based on the technique described by Kumar et al (Kumar, Tan & Steinbach 2005). Once the interviews had been conducted and the transcripts had been prepared, the usable material, by themes, was drawn out from the transcripts through a process called Coding. The coded research material was copied and pasted into separate Microsoft Word files, one for each theme. These files provided an easy look through while writing the Empirical Materials chapter. The qualitative data was then analyzed using the interpretive approach (Miles & Huberman 1994).

The material collected through qualitative methods is invariably unstructured and unwieldy (Bryman & Bell 2007). It is the rough material collected from the field, if the form of videotapes, conversations etc., that form the basis of analysis (Bogdan & Biklen 1992). Due to its complicated nature, there is no standardized approach to the analysis of the qualitative data (Saunders, Lewis & Thornhill 1998).

The raw material resulting from the data gathering process in qualitative research is usually in the form of words, and there are different strategies to deal with words. Miles et al (Miles & Huberman 1994) outlined three approaches for analyzing qualitative data – interpretive, collaborative social research, and social anthropology. Creswell (Crewell 2003) further identified five approaches – case study, biography, phenomenology, grounded theory, and ethnography.

This report uses the interpretive approach to analyze the data as it is used to present a holistic view of data rather than a condensed view. The results of this analysis are discussed in the Analysis section.

RESEARCH LIMITATIONS

The exploratory nature of the research and the majority of experienced practitioners living in distant locations - most successful social media marketing firms are either based in expertise were limited. However, much consideration was given in selecting the suitable interviewees mainly via professional recommendations in the social media industry and a review of their work. Even though care was taken while identifying the right interviewees, the number of interviews meant that, the findings of the research are not tested to be statistically significant. Further to this, mitigating any kind of bias that interviewees might have had, as argued by Robson (Robson 1993), is limited by my understating of the subject area and interpretation.

ETHICAL CONSIDERATIONS

The interviewees have been informed of the academic purpose of this study. They have granted permission for using their details in the report. The research is based on analyzing primary and secondary data using frameworks already developed by researchers. A soft copy of the completed report will be made available to the interviewees. This research also has approval from University of Edinburgh and gives rise to no ethical issues.

PERSONAL EXPERIENCE

My interest in social media started as a quest for more information. The new applications on the mobile devices (for example, Flipboard on iPad) have made it much easier to identify, target and receive information from the specific sources that are of interest to me. Twitter has increasingly become an important online source for the latest information in almost any area –including but not limited to news, ideas, opinions and reflections.

In November 2010, I came across a tweet about a product called KIK - a cross platform instant messaging application. It caught my interest as I had seen Blackberry users raving about the Blackberry Messenger (BBM). A simple search for #KIK on Twitter revealed that thousands of people were talking about this product, downloading it and giving positive or negative feedback about it. Two weeks after launching, KIK was downloaded on more than 1 million mobile devices. Achieving such a huge number of downloads in such a short span of time caught my interest.

A few questions sprung to mind straight away. How did they manage to do it? Was it just luck or a carefully carved strategy? Are the marketers aware of the ongoing conversation on the social networks about their brand, and if so, how do they manage it? Then the conversation on Twitter slowed down considerably, and the question I asked myself was what could marketers do to sustain the consumer interest in their brand?

In December 2010, I came across another app, Viber, a competitor to Skype, providing free calls between iPhones over a 3G network. It was a similar story with Viber in that it achieved 1 million download in just 5 days. KIK now did not seem like a one off app that got lucky. I started looking for already existing research on these topics. It was surprising when I found that most of the academic material was quite dated, more so because social media and its usage is changing so rapidly at the moment that it is difficult for the academic world to catch up.

I also realized that Twitter was one of the main sources for finding information. Most opinions leaders, industry specialists and even industry and academic journals tweet about their papers, blog entries, comments on other members' articles and ideas. However, Twitter is like a stream of consciousness, with ideas and thoughts flowing vertically on a 19-inch screen - just like credits flowing on the screen at the end of a movie, but only much faster. This stream of data is on a variety of subjects and comes from hundreds of people you follow on Twitter. Once I managed to get a handle on the conversations that were specific to social media marketing, I started reading blogs, industry and academic journals and whitepapers to understand how companies go about using social media as an effective marketing tool.

This gave me deep insights into: how organizations go about using social media, latest development and opinions from the thought leaders in social media, reviews and opinions on various strategies and tools used in social media marketing, and importance of sentiment analysis, or Netnography (Kozinets 2010) of the consumer conversation about a certain brand. This secondary research also allowed me to get an understanding of the key areas surrounding viral marketing of the products. It helped me develop a framework for the interviews. Twitter also helped me identify some of the key social media influencers and converse with them. For example, connection with the Online Community Manager of KIK was established using Twitter, and after a few tweets back and forth she agreed to participate in my research.

CONCLUSIONS

This chapter contains the summary of the results described in Chapter 5 with particular reference to the research questions. It also contains my personal reflections on the research and the process of getting to the results. Furthermore, contributions of this research to existing literature, lessons for the business are discussed before highlighting the future research opportunities.

CONTRIBUTION TO EXISTING KNOWLEDGE

The results of the research provide insights into how marketers perceive viral marketing and social media marketing in the real world. Following areas stand out in particular:

- Viral messages and viral marketing campaigns for digital products are subtly different in the way of using viral marketing as part of the overall marketing strategy with specific objectives in mind

- Even though there is research and theoretical models for understanding, creating and executing viral marketing campaigns, in real life the tools available to marketers aren't sophisticated enough to manage and monitor all aspects of such campaigns. Organizations are continuously struggling with resource availability for any kind of social media marketing, unless it is well thought out at the business strategy level.

- The outcome of the interviews provides insights into how companies are using social media and viral marketing concepts in real life to market their products and services. It also provides insights into the importance of overall business and marketing strategy while considering viral marketing as an option to market products.

- The research is relevant to marketers and organizations planning to use viral marketing concepts to market their digital products using social media.

FURTHER RESEARCH

Due to the lack of time and resource availability, the number of interviews in the primary research was kept to just four. Further interviews with companies or professional social media consultants and marketers - who either have used viral marketing to market their digital products on social media, or are engaged in the viral marketing process as strategists, policy makers or influencers - would add more authenticity to the results, or clarify them further. Mobile application being one of my interest area may have meant that I had a personal bias in this area while performing the research and articulating the findings. Further research with fresh perspective on other types of digital products like website, online services etc. would further refine the results. Sentiment analysis (netnography) is an emerging field. There are many challenges yet to be overcome, especially when it comes to automatically collecting and performing an accurate sentiment analysis on the conversations that relate to a brand or product on a variety of social networks. Once sentiment analysis becomes a mature field, further research in this area should provide ways to better monitor and manage negative feedback during marketing campaigns.

REFERENCES

1. Subramani, Mani R., & Rajagopalan, Balaji. (2003). Knowledge sharing and influence in online social networks via viral Marketing. ***Communications of the ACM,*** 46(12), 300-307.

2. Angel, Robert, & Sexsmith, Joseph. (2009, July-August). Social networking: the view from the C-suite. ***Ivey Business Journal***, 14818248, 73(4).

3. Baker, Bill. (2009). Your customer is talking - to everyone; Social media is the new channel for Customer connection. ***Information management,*** 19(4), 1-4. New York.

4. Smith, Tom. (2009). The social media revolution. ***International Journal of Market Research***, 51(4), 559-561.

5. Mullen, Megan Gwynne. (2003). A Scheduling and Programming Innovator. ***The Rise of Cable Programming in the United States: Revolution or Evolution?*** (1st Ed.). ISBN 0-292-75273-3, 170, Austin, TX: University of Texas Press.

6. Howard, Theresa. (2005, June 23). USAToday: Viral advertising spreads through marketing plans. ***USA Today.*** Retrieved on May 27, 2010.

7. Rayport, Jeffrey. (1996, December 31). ***The Virus of Marketing***. Fast Company. Retrieved on 6 October, 2012.

8. Subramani, R., & Rajagopalan, B. (2003). Knowledge Sharing and Influence in Online Social Networks via Viral Marketing. ***Communications of the ACM,*** 8(12), 300-307.

9. (2012, July 30). ***Viral Marketing*** (Night & Day Graphics). Retrieved on 6 October, 2012.

10. (2012, September 11). ***Viral Marketing – Understanding the Latest Catchword*** [Video Marketing Bot Pro.]. Retrieved on 6 October, 2012.

*****

Editor-In-Chief
Pezzottaite Journals,
24, Saraswati Lane, Bohri, Near Modern Dewan Beverages,
Jammu Tawi – 180002, Jammu and Kashmir, India.
(Mobile): +91-09419216270 – 71
editorinchief@pezzottaitejournals.net,contactus@pezzottaitejournals.net

AN ANALYSIS OF E-COMMERCE WEBSITE DESIGN FEATURES THAT IMPACT ONLINE PURCHASING OF INDIAN CONSUMER USING MULTIVARIATE FACTOR ANALYSIS

Ritesh Kumar Srivastava[13]

ABSTRACT

Now a days, online shopping is been mushrooming among every type of consumer worldwide. Thus, it is an important for companies to be aware in order to know how to gain benefits of this electronic channel. The aim of this research is to identify and explore the key factors that impacts customer preference of traversing websites and availing its services in India. Seven key factors of consumer intention of online shopping have been identified.

The results revealed that main features a consumer appreciates are 'Ease of use' and 'Content availability with its nice presentation" about the products and services. These features are correlated with people attitude directly to do online shopping. The analytical results of the collected data through survey showed that there is a highly significant relationship of website design features upon retention of inline purchasers.

KEYWORDS

E-Commerce Strategy, Online Consumer Behavior, Indian Web Market, Website Design, Customer Loyalty etc.

INTRODUCTION

There are various features of e-commerce website through which customers are attracted to online shopping. It is not only because of the high level of convenience, but also because of the wide range, competitive pricing and greater access to information. Business organizations currently adopting e-commerce channels to offer greater shopping experience to consumer because it is much lower cost compared to bricks and mortar stores. It is ubiquitous in nature, offers global reach and follows universal standards. Richness and interactivity made e-commerce being adopted worldwide by the companies and information density and customization factors attract customers for better shopping experience.

Designers of e-commerce websites are concerned with the effects of information density provided to consumers in various forms. A successful e-commerce website is not just a good-looking website with dynamic technical features but to disseminate information, it is about building relationships and generating revenue.

User-centered design is an important matter. Expectations and needs are not same for everyone. Their demographic characteristics viz. Age, gender, experience, culture is all primary factors is a concern. To satisfy each customer's expectation and identifying the customer's wants and needs gives the customer a reason to retain and generate desire to shop. It is important that the website communicate to the customer that the company cares about them. Users with regular online purchasing experience focus more on the comfort ability, while novice users focus on information density. Various researches on e-commerce website have found that the organization of information on the website is important for novice and first-time users, but it gradually declines its importance over time. Information density and content becomes the major factor fetching further visits from the users (Davern, et al, 2001).

Pulling a consumer to use a company's website is a marketing activity, but once a consumer is on website, the sales process starts. This means that whatever they find at the website will ultimately determine whether they will make a purchase or not. Actually, an e-commerce web page and website are first, a software interface, which assists in decision making of consumer. More and more websites using interactive consumer-decision support systems to help shopper make choices. Recommendation agents are programs that can suggests a product based on either consumer surveys or a review of a consumer's profile.

The question is; what makes for an effective user interface design that consumer perceive to be useful and easy to use. In an exploratory study of website credibility, the top three factors in website credibility were design look, information design/structure, and information focus (Fogg et al, 2002).

In this research seven basic design features are identified that were necessary from a business point of view, to attract and retain customers. The website must be functional, informative, employ simple navigation (ease of use), use redundant navigation, and make it easy for customers to purchase and feature multi-browser functionality, simple graphics, and legible text. Researchers have also found a number of other design factors that website designers must be aware off.

Websites that offer a persuasive experience in the sense of providing entertainment with e-commerce or interactivity, or that are perceived as fun to use, are more successful in fetching and retaining visitors. Websites with more information and content density increases probability of purchasing a product or service, while simplicity of design is hard to define. eMarketer reports that more than fifty percent shopping carts are neglected by confused and upset customers (eMarketer, Inc, 2007g).

[13]Assistant Professor, Department of MBA, Sherwood College of Engineering Research and Technology, Uttar Pradesh, India, sri.rit@gmail.com

Responsiveness of a website is also important to credibility. A study of large firms website found that while 90% of firms have a privacy policy, 75% don't tell users how to have their personal information destroyed so it can't be used by others, 12% of large companies didn't respond to e-mail queries, and 21% responded only to half the queries sent to their websites. In general, large companies with websites receive favorable respect ratings for "simplicity of design and use" but weak ratings on responding to customers (walker, 2004; Customer Respect Group, 2005).

No matter how successful the offline and online marketing campaign, a website that fails to deliver information, customer convenience, and responsiveness spells disaster. Attention to these website design features will help ensure success.

OBJECTIVES OF STUDY

Objective of this study is to analyze and decide the priority factors that are involved in website design and its features, which affect online purchasing. There are seven basic design features identified to be tested which offers a necessary framework from a business point of view, to attract and retain customers while making a transaction and shopping online. They are:

1. Fast downloading time,
2. Few clicks to purchase,
3. Content/ Information,
4. Responsiveness,
5. Compelling experience,
6. Customer decision support tools,
7. Easy product navigation.

RESEARCH METHODOLOGY

The study was carried out 2012 to understand and analyze the design features of e-commerce websites in India. A sample of 160 online consumer respondents was drawn from various part of the country. The respondents were asked to state their level of agreement and disagreement of above mentioned factors on the 5-point scale, where: 5 = Strongly Agree, 4 = Agree, 3 = Neutral, 2 = Disagree and 1 = Strongly Disagree. Convenient sampling method was adopted for administering the questionnaires.

Methods of Data Collection

Primary data was collected with the help of structured questionnaire circulated among 200 regular online consumers across India from which 160 completed questionnaire respondents were selected. Study was carried out in the month of November 2012. Factor Analysis was carried out by using SPSS. Factor analysis test was advocated to find out various priority factors among the various factors important to attract and retain online customers.

Areas of Data Collection

The data has been collected from regular online consumers purchasing from various part of the country.

DATA ANALYSIS AND INTERPRETATION

In order to establish the strength of the factor analysis solution, it is essential to establish the reliability and validity of the obtained reduction. This is done with KMO and Bartlett's test of sphericity. The result on test is given below:

Table-1: KMO and Bartlett's Test

Bartlett's Test of Sphericity	Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.620
	Approx. Chi-Square	120.870
	Df	21
	Sig.	.000

Sources: Authors Compilation

It may be noted that the value of KMO statistics is 0.620 which greater than 0.5, which indicates that factor analysis could be used for the given set of data. Further p value is 0.00 which is less than 0.05 indicates the rejection of the hypothesis that the correlation matrix of the variables is insignificant.

It is important to indicate how much of each variable is accounted for by the underlying factors taken together. It is a measure of the percentage of the variable's variation that is explained by the factors. The communality of each variable is explained in below table.

Table-2: Communalities

	Component	
	1	**2**
Fast downloading time	1.000	.425
Few clicks to purchase	1.000	.571
Content/ Information	1.000	.637
Responsiveness	1.000	.174
Compelling experience	1.000	.691
Customer decision support tools	1.000	.231
Easy product navigation	1.000	.566

Sources: Authors Compilation
Extraction Method: Principal Component Analysis

Table-3: Total Variance Explained

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	2.024	28.908	28.908	2.024	28.908	28.908
2	1.272	18.169	47.077	1.272	18.169	47.077
3	.989	14.126	61.203			
4	.956	13.659	74.862			
5	.755	10.785	85.647			
6	.544	7.770	93.417			
7	.461	6.583	100.000			

Sources: Authors Compilation
Extraction Method: Principal Component Analysis

We note from the above table that there are two factors with eigen value more than one.

Table-4: Component Matrix[a]

	Component	
	1	**2**
Fast downloading time	-0.578	-0.3
Few clicks to purchase	-0.629	0.719
Content/ Information	0.739	0.209
Responsiveness	-0.062	0.412
Compelling experience	0.803	-0.216
Customer decision support tools	0.306	0.371
Easy product navigation	-0.063	0.75

Sources: Authors Compilation
a. Extraction Method: Principal Component Analysis.
2 components extracted.

The above-mentioned two factors should be statistically independent. This means that the correlation between the two factors scores should be zero.

The next task is to interpret the factor-loading matrix called the component matrix. In order to do so and to be able to interpret results in a better way a factor rotation matrix is calculated using Varimax rotation option in SPSS, which results in independent factors. The result obtained after Varimax rotation is given below:

Table-5: Rotated Component Matrix[a]

	Component	
	1	**2**
Fast downloading time	-0.64	-0.123
Few clicks to purchase	-0.483	0.701
Content/ Information	0.794	0.077
Responsiveness	0.058	0.49646
Compelling experience	0.708	-0.436
Customer decision support tools	0.4	0.368
Easy product navigation	0.154	0.737
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.		
a. Rotation converged in 3 iterations.		

Sources: Authors Compilation

FINDINGS

The cutoff point is decided to 0.7. The two variables corresponding to factor 1 having a factor loading above 0.7 are 'Content/ Information' and 'Compelling experience' to purchase. The variables corresponding to factor 2 for which the factor loading is greater than 0.7 is 'Few clicks to purchase' and 'Easy product navigation'.

CONCLUSIONS

In this research seven basic design features are identified that were necessary in order to attract new online customers and retain regular online customers. The website must be purposeful, informative, with easy navigation features, and make it easy for customers to purchase with simple graphics, and clear text. Researchers have found a number of other design factors that website designers must be aware off. In our conclusion after factor analysis, it is observed that the factor1 comprising of 'Content/ Information' and 'Compelling experience' can be termed as Information density and its presentation and factor2 comprising of the variables 'Few clicks to purchase' and 'Easy product navigation' can be termed as User Comfortability. This shows that most important factor which makes user to become loyal user and encourage purchase among online consumer in India is Information density and its presentation followed by User comfortability given by website to the online user. A novice user always desire notable information presented in user friendly way and regular user seeks of comfortability and shorter cycles of purchase.

REFERENCES

1. Ganguly, B., Dash, S. B., Cyr, D., & Head, M. (2010). The effects of website design on purchase intention in online shopping: the mediating role of trust and the moderating role of culture. ***Int. J. Electronic Business*****,** 8(4/5), 302-330.

2. Chen, Q., & Wells, D. W. (1999, September/October). Attitude toward the site. ***Journal of Advertising Research***. 39, 27-37.

3. Akram, Jalal Karim. (2011). Evaluating the influence of reliability, usability, usefulness and website design on the satisfaction of online consumers. ***Research Journal of Economics, Business and ICT****,* 2, 29-32.

4. Bigne, Enrique. (2005). The impact of internet user shopping patterns and demographics on consumer mobile buying behavior. ***Journal of Electronic Commerce Research***, 6.

5. Lee, Y., & Kozar, A. K. (2008). Designing usable online stores: A landscape preference perspective. ***Inf. Manage***, doi:10.1016/j.im.2008.11.002

6. Peterson, R. A., Balasubramanian, S., & Bronnenberg, B. J. (1997). Exploring the implications of the Internet for consumer marketing. ***Journal of the Academy of Marketing Science***, 25, 329-346.

7. Cyr, D. (2008). Modeling website design across cultures: relationships to trust, satisfaction and e-loyalty. ***Journal of Management Information Systems,*** 24(4), 47-72.

8. Fogg, B. J. (2002). ***Persuasive Technology: Using Computers to Change What We Think and Do*****.** Morgan Kaufmann.

9. Laudon, C. K., & Traver, G. C. (2011). ***E-commerce, Business Technology Society*** (4th ed.). Noida: Dorling Kindersely (India) Private Limited.

10. Ethier, J., Hadaya, P., Talbot, J., & Cadieux, J. (2006). B2C web site quality and emotions during online shopping episodes: An empirical study. ***Information and Management***, 43, 627–639.

11. Fogg (et al.). How Do Users Evaluate Credibility of Websites? ***Consumer Reports Web watch.*** Retrieved from www.consumerwebwatch.org

12. Kim, J., Lee, J., Han, K. & Lee, M. (2002). Business as buildings: metrics for the architectural quality of internet businesses. ***Information Systems Research,*** 13, 239–254.

13. Venkatesh, V. & Agarwal, R. (2006). Turning visitors into customers: a usability-centered perspective on purchase behavior in electronic channels. ***Management Science*** 52(3), 367–382.

14. Retrieved from http://en.wikipedia.org/w/index.php?title=Online_shopping&oldid=536847751

15. Retrieved from http://www.bonsoni.com/blog/research-shows-word-of-mouth-drives-online-sales/

16. Retrieved on February 10, 2010 from http://www.ecommercetimes.com/story/69105.html.

17. Retrieved on February 4, 2012 from http://www.consumerwebwatch.org/pdfs/stanfordPTL.pdf

*****

SOME REFLECTION ON GRAMMEEN CREDIT WITH REFERENCE TO GRAMMEEN BANK BANGLADESH

Gaurangkumar C. Barot[14]

ABSTRACT

Whenever I use the word, "Micro credit" I actually have in mind Grameen type microcredit or Grameen credit. However, if the person I am talking to understands it as some other category of microcredit my arguments will not make any sense to him. Let me list below the distinguishing features of Grameen credit. This is an exhaustive list of such features. Not every Grameen type programme has all these features present in the programme. Some programmes are strong in some of the features, while others are strong in some other features. Overall, they display a general convergence to some basic features based on which they introduce themselves as Grameen replication programmes or Grameen type programmes.

KEYWORDS

Grammeen, Credit, Bangladesh, Women, Finance etc.

INTRODUCTION

The origin of Grameen Bank can be traced back to 1976 when Professor Muhammad Yunus, Head of the Rural Economics Program at the University of Chittagong, launched an action research project to examine the possibility of designing a credit delivery system to provide banking services targeted at the rural poor. The Grameen Bank Project (Grameen means "rural" or "village" in Bangla language) came into operation. Grameen Bank (GB) has reversed conventional banking practice by removing the need for collateral and created a banking system based on mutual trust, accountability, participation and creativity.

GB provides credit to the poorest of the poor in rural Bangladesh, without any collateral. At GB, credit is a cost effective weapon to fight poverty and it serves as a catalyst in the overall development of socio-economic conditions of the poor who have been kept outside the banking orbit on the ground that they are poor and hence not bankable. Professor Muhammad Yunus, the founder of "Grameen Bank" and its Managing Director, reasoned that if financial resources can be made available to the poor people on terms and conditions that are appropriate and reasonable, "these millions of small people with their millions of small pursuits can add up to create the biggest development wonder." As of October 2011, it has 8.349 million borrowers, 97 percent of whom are women.

With 2,565 branches, GB provides services in 81,379 villages, covering more than 97 percent of the total villages in Bangladesh. Grameen Bank's positive impact on its poor and formerly poor borrowers has been documented in many independent studies carried out by external agencies including the World Bank, the International Food Research Policy Institute (IFPRI) and the Bangladesh Institute of Development Studies (BIDS).

OBJECTIVES OF STUDY

- Grammeen bank's loan distributes activity in Bangladesh.
- Grammeen bank's Deposits performance through the members in Bangladesh.
- Grammeen bank's profit performance in Bangladesh.

RESEARCH METHODOLOGY

Data Source

Secondary data was collected from various Grammeen bank's annual report, Books, Journals, Published Reports, Newspapers, Magazines and Electronic Data.

Research Design

For the study purpose, Grammeen credit institute is selected i.e. Grameen bank, Bangladesh. Last five year statistic data is collected. Hence, through the secondary data the hypothesis is formulated and analyzed with the effective test. The Comparative study (Intra-comparison) of recent five (05) years is taken.

Table wise interpretation is given for the betterment of the different activities. For the hypothesis study purpose various Statistics tools is used such as t- test, Trend Analysis, Mean, Mode, Medium, Co-relation, charts, Comparative Tables study etc. All tests calculated and direct result mention here in this study.

[14]Assistance Professor, Department of Commerce, Silvassa Institute of Higher Learning, A Government Aided College (Arts Commerce and Science), Silvassa (Naroli) D & N. H. (U.T.) India, gaurangcbarot@gmail.com

Hypotheses of Study

Case-I

- H_0 = Grameen bank significantly distributes loan.
  H_1 = Grameen bank does no significantly distribute loan.

Case-II

- H_0 = Grameen Bank significantly received Deposits from the members.
  H_1 = Grameen Bank did not significantly receives Deposits from the members.

Case-III

- H_0 = Grameen bank earned significant profit.
  H_1 = Grameen bank did not earned significant profit.

DATA ANALYSIS AND INTERPRETATION

Case-I: Hypotheses Testing

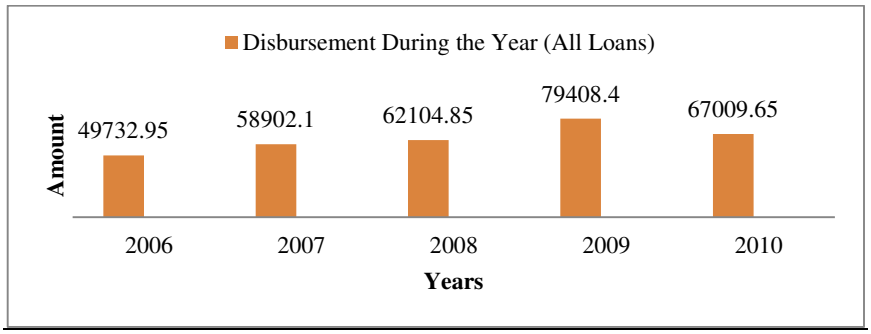
- H_0 = Grameen bank significantly distributes loan.
- H_1 = Grameen bank does no significantly distribute loan.

Table-1: Disbursement during Year (All Loans)
(Amount in Million Taka)

Performance Indicator	2006	2007	2008	2009	2010
Disbursement During the Year (All Loans)	49732.95	58902.10	62104.85	79408.40	67009.65

Sources: Various Annual Report of Grameen bank Bangladesh

Graph-1: Disbursement during the Year (All Loans)



Sources: Graphical Representation

Hypotheses Testing

Here, Observation is less then < 30, hence, t- test is recommended.

For, t – test μ value is necessary. In this case, Mode (Z) value is recommended for μ value,

$H_0 = \mu = 62104.85$, $H_1 = \mu \neq 62104.85$,

$\sum xi = 317157.95$, $\sum di = 0$, $\sum di^2 = 477990227.7$, $\bar{x} = 63431.59$

For, t -test standard deviation (S) is = 9777.4253

Calculated t -test value is = 0.27138842

Degree of freedom (d.f.) = 4 and Table value at 5% level of significant = 2.776

As calculated value of 't' is less then the table value, Hence, H_0 = is Accepted, $H_0 = \mu = 62104.85$.

Interpretation

Hypothesis result shows that Grameen bank significantly distributes all types of loan. Activities performance growth rate was increased in positive manner. It was in the year 2006 contributed 49732.95 Cr. Amt. If I consider this performance as base performance than in the 2007 year, performance was 58902.1Cr. of Amt. (18.43 % growth rate increased then the base year) was

achieved. In The 2008 year, performance was 62104.85Cr. of Amt. (24.87 % growth rate increased then the base year) was achieved. In The 2009 year, performance was 79408.4Cr. of Amt. (59.66 % growth rate increased then the base year) was achieved. In the 2010 year, performance was 67009.65 Cr. of Amt. (34.73 % growth rate increased then the base year) was achieved.

Case-II: Hypotheses Testing

- H_0 = Grameen Bank significantly received Deposits from the members.
  H_1 = Grameen Bank did not significantly receives Deposits from the members.

Table-2: Deposits of Grameen Bank Members (Balance)
(Amount in Million Taka)

Performance Indicator	2006	2007	2008	2009	2010
Deposits of GB Members (Balance)	27298.19	29533.51	34923.62	44823.49	51236.25

Sources: Various Annual Report of Grameen bank Bangladesh

Graph-2; Deposits of GB Members (Balance)

Sources: Graphical Representation

Hypotheses Testing

Mean(x) = 37563.012 Amt. **Median (M)** = 34923.62 Amt. **Mode (Z) =** 51236.25 Amt.

Here, Observation is less then < 30, hence, t- test is recommended.

For, t – test μ value is necessary. In this case, Mode (Z) value is recommended for μ value,

$H_0 = \mu = 51236.25$, $H_1 = \mu \neq 51236.25$,

$\sum xi = 187815.06$, $\sum di = -1.45$, $\sum di^2 = 416477841.4$, $\bar{x} = 37563.012$

For, t -test standard deviation (S) is = 9126.640

Calculated t -test value is = 0.000328306

Degree of freedom (d.f.) = 4 and Table value at 5% level of significant = 2.776

As calculated value of 't' is less then the table value, Hence, H_0 = is accepted, $H_0 = \mu = 51236.25$

Interpretation

Hypothesis result shows that Grameen Bank significantly received Deposits from the members. Activities performance growth rate was increased in positive manner. It was in the year 2006 contributed 27298.19 Cr. rupees If I consider this performance as base performance than in the 2007 year, performance was 29533.51 Cr. of Amt. (08.18% growth rate increased then the base year) was achieved. In The 2008 year, performance was 34923.62 Cr. of Amt. (27.93 % growth rate increased then the base year) was achieved. In The 2009 year, performance was 44823.49 Cr. of Amt. (64.19 % growth rate increased then the base year) was achieved. In the 2010 year, performance was 51236.25 Cr. of Amt. (87.69 % growth rate increased then the base year) was achieved.

Case-III: Hypotheses Testing

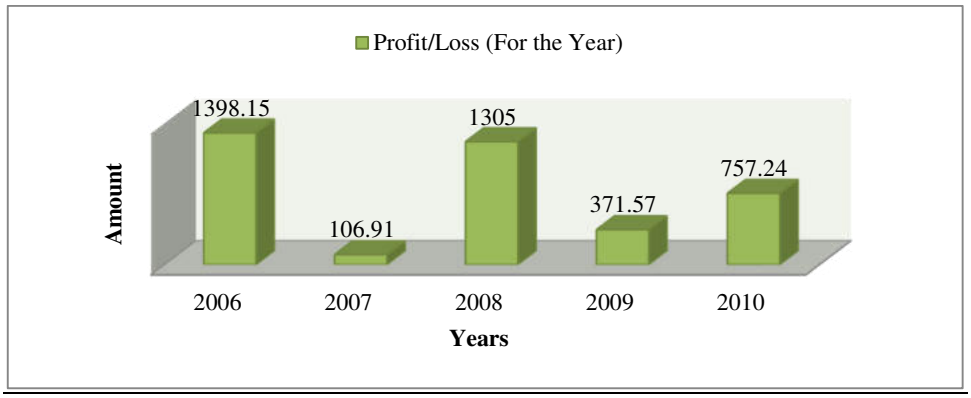
- H_0 = Grameen bank earned significant profit.
  H_1 = Grameen bank did not earned significant profit.

Table-3: Profit (For the Year)
(Amount in Million Taka)

Performance Indicator	2006	2007	2008	2009	2010
Profit/Loss (For the Year)	1398.15	106.91	1305.00	371.57	757.24

Sources: Various Annual Report of Grameen bank Bangladesh

Graph-3: Profit (For the Year)



Sources: Graphical Representation

Hypotheses Testing

Mean(x) = 787.774 Amt. **Median (M)** = 1305.00. **Mode (Z) =** 1398.15

Here, Observation is less then < 30, hence, t- test is recommended.

For, t – test μ value is necessary. In this case, Mode (Z) value is recommended for μ value,

$H_0 = \mu = 1398.15$, $H_1 = \mu \neq 1398.15$,

$\sum xi = 3938.87$, $\sum di = 5.684$, $\sum di^2 = 1277815$, $\bar{x} = 787.774$

For, t -test standard deviation (S) is = 505.531

Calculated t -test value is = 1.207395788

Degree of freedom (d.f.) = 4 and Table value at 5% level of significant = 2.776

As calculated value of 't' is less then the table value, Hence, H_0 = is accepted, $H_0 = \mu = 1398.15$

Interpretation

Hypothesis result shows that Grameen bank earned significant profit. Performance growth rate was increased in positive manner. It was in the year 2006 contributed 1398.15 Cr. Amt. If I consider this performance as base performance than in the 2007 year, performance was 106.91 Cr. of Amt. (90 % growth rate decreased then the base year) was achieved. In the 2008 year, performance was 1305 Cr. of Amt. (10 % growth rate decreased then the base year) was achieved. In the 2009 year, performance was 371.57 Cr. of Amt. (72 % growth rate decreased then the base year) was achieved. In the 2010 year, performance was 757.24 Cr. of Amt. (40 % growth rate decreased then the base year) was achieved.

CONCLUSIONS

From the first hypothesis of study, it is recommended that the way Grameen bank significantly distributes all types of loan and that should carry in the same manner with increasing growth rate.

From the second hypothesis, result of the study it is recommended that the way Grameen Bank significantly received Deposits from the members and that should carry in the same manner with increasing growth rate.

From the third hypothesis of the study it is recommended that the way Grameen bank earned profit but a lot more to do for the betterment of the incising the profit and should carry in the more action for improve the growth rate.

REFERENCES

1. Armendariz, de Aghion, Beatriz, & Jonathan, Morduch. (2000). Microfinance beyond group lending. ***Economics of Transition,*** 8, 401-420.
2. Barnes, Carolyn, Gayle Morris, & Gary, Gaile. (1999). An Assessment of Client of Microfinance Programs in Uganda. ***International Journal of Economic Development***, 1, 80-121.
3. Cole, Rebel A. (1998). The importance of relationships to the availability of credit. ***Journal of Banking and Finance***, 22, 959-977.
4. Kothari, C. R. (2007). ***Research Methodology***. Delhi: Wishwa Publication.
5. Wydick, W. Bruce. (2002). Microfinance among the Maya: Tracking the progress of borrowers. ***Development and Change***, 33, 489-509.
6. Yunus, Mohhamad. (2008). ***Grammeen Bank of Bangladesh.***
7. Retrieved from http://www.grameenfoundation.org/

*****

A STUDY ON MULTI-STAGE CASCADING CLASSIFICATION FOR DIGITAL MAMMOGRAPHY

V. Gowri Prakash[15] K. R. Valluvan[16]

ABSTRACT

Mammography is the most widely used and accepted non-invasive medical imaging technique used for breast cancer screening. It refers to the imaging of the breast using x-rays and the image obtained is called a mammogram. Deciding on whether a lesion is benign or malignant from a mammogram is a bit tricky but there are definite clues for it. Image segmentation is a very important step towards the identification of affected region (or set of pixels) in a mammogram image of a patient to detect the breast cancer, which calls for the expertise of a physician and / or a radiographer. The mammography is the most effective procedure for an early diagnosis of the breast cancer. The goal of segmentation is to detect abnormalities and to extract the entire suspicious mass region from mammograms. In this paper, we have reviewed different Multi-Stage Cascading Classification of mammograms for the detection of breast cancer. Different classification methods are studied and their merits and drawbacks are outlined.

KEYWORDS

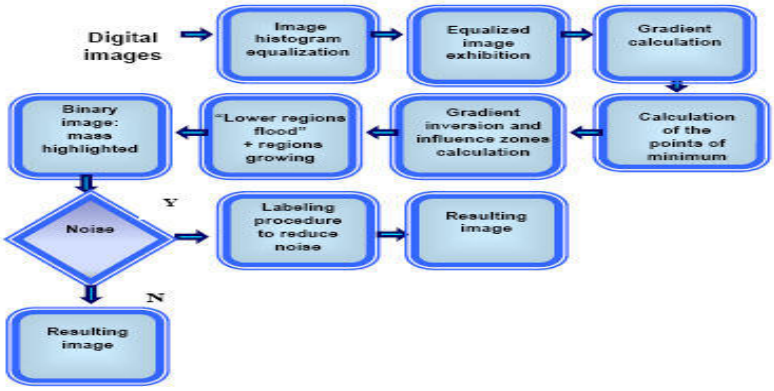
Breast Cancer, Preprocessing, Image Segmentation, Cascading Classification, Digital Mammography etc.

INTRODUCTION

Breast cancer is a malignant tumor arising from breast cells. The malignant tumor can destroy the nearby cells and spread to different part of the body. Breast cancer is one of the most dangerous types of cancer among women all over the world and it can affect men, although the male breast cancer is rare [1]. The reports of recent studies show that breast cancer affects one of every eight women in the United States and one of every ten in Europe [2]. Therefore, the effective way to reduce the mortality is early detection and treatment of breast cancer. Mammography is the most commonly used methods for early detection and diagnosis of breast cancer [3].

Mammography is a specific type of imaging that uses a low-dose x-ray system to examine breasts. It is used to find tumors and to identify the difference between benign and malignant cells. Benign is a noncancerous cell and it do not spread to other parts of the body. However, malignant is a cancerous cell and that can spread to other parts of the body [4]. It has been found that the interpretation of mammograms by radiologist, many a times give high rates of false positive cases [5]. In the breast cancer, analysis the estimated sensitivity of radiologist is only about 75%. Therefore, different effective and efficient diagnostic methods based on image processing algorithm have been developed by the researches. Different segmentation methods have been suggested to improve the accuracy of interpretation. The below figure shows the preprocessing of the mammogram.

Figure-1: Flow Chart of the Preprocessing



Sources: Mammography Screening: The Nordic Cochrane Centre, Rigshospitalet

Segmentation methods are based on some pixel or region similarity measure in relation to their local neighborhood. These similarity measures in texture segmentation methods use some textural spatial-spectral-temporal features such as Markov random field statistics (MRF) [8-10], co-occurrence matrix based features [11], Gabor features [12], local binary pattern (LBP) [13], autocorrelation features and many others. A number of image processing methods have been proposed to perform this task. S. M.

[15]Assistant Professor- GR1, Department of Computer Science & Engineering, S. N. S. College of Engineering, Tamil Nadu, India, gowriprakash83@gmail.com
[16] Professor & Head, Department of ECE, Velalar College of Engineering & Technology, Tamil Nadu, India, krvaluvan@yahoo.co.in

Li et al. [14] and W. Qian et al. [15] have proposed using modified and weighted median filtering, respectively, to enhance the digitized image prior to object identification. D. Brzakovic et al. [16] used threading and fuzzy pyramid linking for mass localization and classification. Other investigators have proposed using the asymmetry between the right and left breast images to determine possible mass locations. Yin ***et*** *al.* [17] uses both linear and nonlinear bilateral subtractions while the method by Lau *et al.* [18]. Relies on "structural asymmetry" between the two breast images. Recently Kegelmeyer has reported promising results for detecting speculated lessons based on local edge characteristics and Laws texture features [19-21].The above methods produced a true positive detection rate of approximately 90%. Various segmentation techniques have been proposed based on statistically measurable features in the image [22-27] Clustering algorithms, such as k-means and ISODATA operate in an unsupervised mode and have been applied to a wide range of classification problems.

Techniques Used for Breast Cancer Measurement

Breast imaging analysis can be performed using X-rays, magnetic resonance, nuclear medicine or ultrasound [1].

X-ray Mammography

X-ray Mammography is commonly used in clinical practice for diagnostic and screening purposes [2]. Mammography provides high sensitivity of fatty breast and excellent demonstration of micro calcifications; it is highly indicative of an early malignancy.

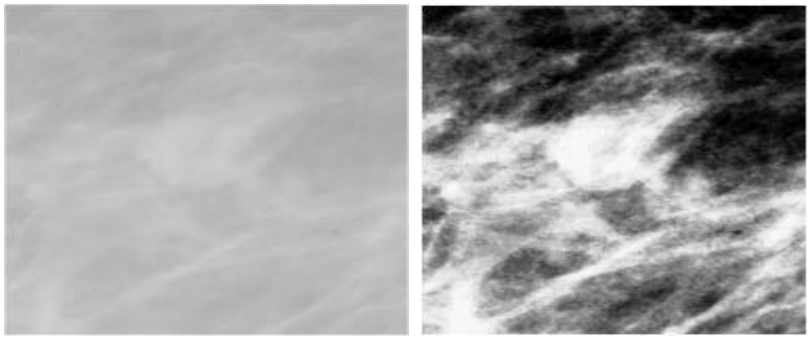
MRI of the Breast

Magnetic Resonance Imaging is the most attractive alternative to Mammography for detecting some cancers, which could be missed by mammography. In addition, MRI can help radiologists and other specialists determine how to treat breast cancer patients by identifying the stage of the disease

PREPROCESSING

Images corresponding to high-density breasts, which show poor contrast in the Region of Interest (ROI). This step produces images with high contrast for histogram equalization. Figure 2 shows the original region of interest extracted from the mammogram and figure 3 shows the region of interest after preprocessing.

Figure-2 **Figure-3**



Sources: Mini-MIAS Database

FEATURE EXTRACTION

The first step in normal analysis is to extract characteristic features from IE. From the perspective of pattern classification, feature extraction is a very important step in that the ultimate performance of the system is not determined by optimal parameters of the classifier, but by the intrinsic seperatibility of the feature vectors. The characterization of normal tissue poses a challenge due to the complexity of normal tissues and the fact that a normal mammogram is not well-defined [7]. Subtle breast cancers may not be easily distinguished from the surrounding normal tissue. The heterogeneous nature of different breast cancers of different sizes also poses challenges for feature extraction. In this section, we will describe several feature sets that we believe can be used to separate normal and abnormal regions. All of our features are extracted from 512£512 regions of IE. There are four types of features extracted from each region.

Curvilinear Features

Though normal breast tissue may have a very different appearance, unequivocally normal breast areas are characterized by curvilinear markings. These curvilinear structures are the distal structures of the breast tissue, which are lactation pathways [7]. The curvilinear markings are not randomly oriented, but rather tend to radiate from the nipple toward the chest wall. If a tumor or a scar appears in an area, the surrounding curvilinear structure is disturbed, usually appearing as a random pattern or is partially absent. Curvilinear structures have been extensively studied for characterizing normal breast tissues.

Gray Level Co-occurrence Features

Texture represents the spatial arrangement of the pixels in a region. Characterization of spatial patterns can be adequately specified by a 2D spatial dependence matrix known as the Gray Level Co-occurrence Matrix (GLCM). Each entry (*I; j*) of the matrix in row *I* and column *j* is the relative frequency of occurrence of pair wise gray levels *i; j* separated by a distance *d* and a direction. The distance *d* can be adjusted to match the size of basic texture elements in the image. The direction parameter can be specified in the direction of the spatial repetition period of the basic texture element.

Gabor Features

2D Gabor filters describe the 2D receptive field profiles of simple cells found in the visual cortex of vertebrate animals. They are consistent with the human vision system (HSV). Gabor filters have been successfully used in describing texture information. A Gabor filter has as its impulse response a Gaussian modulated sinusoidal plane wave:

$$g(x,y) = \frac{1}{2\pi\,\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right] + 2\pi j w_x$$

Where *w* is the modulation frequency, *x*, *y* are the coordinates in the spatial domain, and σ_x and σ_y are the standard deviations in the *x* and *y* direction.

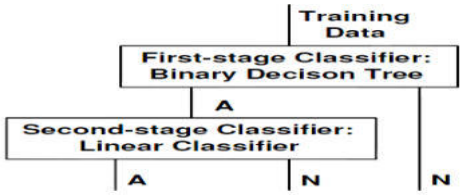
METHODS OF MULTI-STAGE CASCADING CLASSIFICATION

In the previous section, we defined features characterizing mammogram regions. Our next step will be to classify each mammogram region either as normal or as abnormal with a very high correct detection rate. Due to the heterogeneous natures of normal and abnormal classes, a general classification approach may not be adequate to capture each pattern. An efficient classification scheme has to be developed for this purpose. In this session, we will propose a unique multistage cascading classifier to improve the classification performance. In general, we have a training set of feature vectors and the known *class labels* with the objective being to train the classifier such that the classifier can be used to classify new data not contained in the training data. This is known as supervised pattern classification. In some cases, training data having known class labels are not available. The goal is to unravel the underlying similarities and cluster (group) "similar" data together. This is known as unsupervised pattern classification or clustering. Since class labels in the training set provide *a priori* information, supervised classification usually yields better performance than clustering. We will focus on supervised classification with a training data set having known class labels, i.e. the ground-truth information whether a region is normal or abnormal. In practical supervised classification, *N*-dimensional feature vectors *Xi*'s and the corresponding labels *Yi*'s are given as pair entries in the training database, (*Xi*, *Yi*), *i* =1*;* 2*; M*, where *M* is the total number of samples, which is divided into the number of normal samples and the number of abnormal samples i.e. *MS* (*MS+MN = M*). The goal is to find a classification algorithm that can be trained to map *Xi* to*Yi*, and can be generalized for unknown test vectors *Xit*. The trained classifier *GM* from a classification or mapping algorithm, where (*XNi ; Y Ni*) are normal training data, and (*XSi ; Y Si*) are abnormal training data, and the misclassification costs *CN;S* and *CS;N* with *CN;S À CS;N*. There are many different supervised classification algorithms, from a simple linear classifier to nonlinear neural networks, such as linear discriminant analysis (LDA), decision trees , Bayesian networks, neural networks, support vector machines (SVMs) and genetic algorithms .

Two-Stage Cascading Classifier

The structure of our two-stage classifier is shown in Figure 4. The emphasis of this combination scheme is that the first classifier should correctly classify nearly every abnormal region while separating as many normal regions as possible. It is designed to be highly cost-sensitive to make misclassification of an abnormal as "normal" very low. This is the rationale that it is unnecessary to further re-classify the data classified as "normal". This strategy not only reduces the complexity of the combined classification system, but also has a fundamental theoretical formulation, which most combination methods, such as the cascade generalization lack. Only the data classified as "abnormal" by the first classifier are input to the second classifier. The classifier combination scheme is simple and is easily extendable into multiple stages with cascading. Though it has a fundamentally different philosophy of combining classifiers, our two-stage cascading classification can be interpreted as a special implementation of the cascade generalized.

Figure-4: The Structure of Two-stage Cascading Classifier



Sources: Authors Compilation
Note: Class labels: A - Abnormal Class and N - Normal Class

Define our meta-classifier =2(*X*): *X* is an input feature vector:

BEGIN
IF (*X* is classified by first classifier *J*1 as *Normal)*
THEN *X* is *Normal*
ELSE, Use second classifiers *J*2 to reclassify *X*
END

The details of our two-stage classification system are described as follows. A binary decision tree classifier is used as the first classifier because it is one of the most powerful classification tools. Misclassification costs can be specified so that almost all of the training cancer regions are correctly classified. The decision tree algorithm we used is a variation of CART. This algorithm randomly divides the training data into two sets, one to build the pure leaf-node tree and the other to prune; then the roles of two sets are changed and the procedure iterates until converging to a smallest pruned sub tree.

The decision tree classifier is based on a hierarchy of multiple features, which make it difficult to evaluate the performance using Receiver Operating Characteristic (ROC) analysis. Therefore, in addition to improving the classification performance, the second classifier can be used for ROC analysis.

With the design of the first classifier, the input data to the second classifier are only the training regions classified as "abnormal" by the decision tree classifier. In our system, the second classifier is a linear discriminant classifier using Fischer's criterion. The linear classifier is augmented with an adaptive feature selection. Unlike a decision tree classifier, we need to explicitly select a subset of the features. A Sequential Floating Search method is one of the best feature selection methods available.

The principles of floating search feature selection are described as follows:

1. Let *F* be the selected feature subset of size *k*, *S* be the set of other features of size *N ¡ k*, where *N* is the total feature number
2. Start with *k*=0, *F* =Ø, and set the stop number of *K*(*K* <= *N*)
3. Add the most significant feature *f*¤ from *S* to the current subset *F*
4. If the added subset is the best subset of size *k* + 1, then let *k* = *k* + 1, *F* = *F* + *ff*¤*g*, *S* = *S ¡ ff*¤*g*, and go to Step 3; otherwise, go to Step 5
5. Conditionally remove the least significant feature *f*+ from the current subset *F* of size *k*
6. If removed subset is the best subset of size *k¡*1 found so far, then let *k* = *k¡*1, *F* = *F ¡ ff*+*g*, *S* = *S* + *ff*+*g*, and go to Step 5; Else return to Step 3
7. Stop after reaching *k* = *K*

Implementation of the Adaptive Sequential Forward Floating Search (ASFFS) in our study is wrapper based [168], with a linear classifier as the built-in classifier. The criterion used to select the feature subset is A_z, the area under the ROC, generated by the built-in classifier. The advantage of the wrapper method is that feature selection is dependent on and consistent with the classification task. Our two-stage cascading classifier system has the classification power of a decision tree and the simplicity of ROC analysis of a linear classifier.

Multi-Stage Cascading

It is easy to extend our two-stage classifier system to multi-stage. Figure 4.3 shows the extended multi-stage cascading classification system. No matter what kind of classification algorithm it uses, the first through (*n ¡* 1)-th stage classifier

Normal Mammogram Region Identification

Figure 4.2 shows our two-stage classification system for identifying normal mammogram regions. A mammogram region is 512 £ 512 pixels. The enhanced image g*IE* of the region is obtained. Four sets of features are then extracted from *IE* as described are then extracted from *IE* in Section 3.3. From each region, a total of 86 features is obtained. These features were then used to train the two-stage cascading classifier. The classifier was trained with 164 ground-truth cancer regions and 296 normal regions. Among the 164 ground-truth cancer regions, 53 are masses, 56 are speculations and 55 are calcifications. The subtlety rating in the DDSM is used to indicate the distinguishability of a cancer by a radiologist. The lowest subtlety rating (rating 1) means the cancer is very subtle and more difficult to identify. The first-stage decision tree classifier was cost-constrained to correctly classify nearly every cancer region. This resulted in a True Positive Fraction (*TPF*) of 0.99 at False.

COMPARISON OF METHODS

The mammogram segmentation techniques are important to separate suspicious areas of masses or micro-calcification from the background texture. To classify each mammogram region either as normal or as abnormal, Micro classification methods are used. The Two-Stage not only reduces the complexity of the combined classification system, but also has a fundamental theoretical formulation. In Multistage cascading the searched features space is reduced, but it increase the true negative fraction value. The normal mammogram region identification helps to maximize the TNF rate. The comparison results of different methods are shown in Table1.

Table-1: Comparison of Cascading Method

METHOD	MERITS	DEMERITS
Two stage cascading classifier	It is simple and easily extendable.	Computational complexity is high.
Multistage cascading	It increase the true negative fraction trough FPF.	It is applicable to small number of stages only.
Normal Mammogram Region Identification	It maximizes the TNF rate.	Time of maximization process is expensive.

CONCLUSIONS

Digital mammography screening approaches can enable early detection of breast cancer, which reduces the mortality. There are many techniques developed for the detection and classification of masses. This paper discusses the commonly used mammogram image cascading approaches for the early detection and diagnosis of breast cancer. Each method discussed in this paper has its own advantages and disadvantages and any methods can be selected based on the application.

REFERENCES

1. Aboul, Ella Hassanien, & Amr, Bader. (2003). A Comparative study on digital mammography: Enhancement algorithms based on Fuzzy Theory. ***International Journal of Studies in Informatics and Control,*** 12(1), 21–31.

2. Jawad, Nagi, Sameem, Abdul Kareem, Farrukh, Nagi, & Syed, Khaleel Ahmed (2010). Automated Breast Profile Segmentation for ROI Detection Using Digital Mammograms. ***IEEE EMBS Conference on Biomedical Engineering & Sciences,*** Kuala Lumpur, Malaysia.

3. Tabar, L., & Dean, P. (2001). ***Teaching atlas of mammography*** (3rd Ed.). New York: Thime.

4. Gar¯nkel, L., Catherind, M., Boring, C., & Heath, C. (1994). ***Changing trends: An overview of breast cancer incidence and mortality***. Cancer, 74(1), pp. 222-227.

5. Parker, S., Tong, T., Bolden, S., & Wingo, P. (1996). Cancer statistics. ***Cancer J. Clin***, 46, 5{27.

6. Doll, R., Payne, P. & Waterhause, J. (1996). ***Cancer incidence in ¯ve continents.*** Geneve, Switzerland: UICC.

7. Gaughan, R. (1998, September/October). New approaches to early detection of breast cancer makes small gains. ***Biophotonics International***, pp. 48-53 .

8. Tabar, L., Fagerberg, G., S. Du®y, Day, N., Gad, A., & Grontoft, O. (1992). Update of the swedish two-country program of mammographic screening for breast cancer. ***The Radiologic Clinics of North America,*** 30, 187-210.

9. D. Kopans, D. (1989). ***Breast Imaging***. Philadelphia: J.B. Lippincott Company.

10. Howard, J. (1987). Using mamography for cancer control: An unrealized potential. ***CA: A Cancer Journal for Clinicians,*** 37, 33-48.

11. (2002). ***Cancer Facts and Figures***. American Cancer Society.

12. Harvey, J., Fajardo, L., & Innis, C. (1993). Previous mammograms in patients with impalpable breast carcinoma: Retrospective vs blinded interpretation. ***American Journal of Radiology***, 161, 1167-1172.

13. Beam, C., Layde, P., & Sullivan, D. (1996). Variability in the interpretation of screening mammograms by US radiologists, findings from a national sample. ***Archives of Internal Medicine,*** 156, 209-213.

14. Elmore, J., Wells, C., Lee, C., Howard, D. & Feinstein, A. (1994). Variability in radiologists' interpretations of mammograms. ***New England Journal of Medicine***, 331, 1493-1499.

15. Bird, R., Wallace, T., & Yankaskas, B. (1992). Analysis of cancers missed at screening mammography. ***Radiology***,184, 613-617.

16. Sickles, E., Ominsky, S., Sollitto, R., Galvin, H., & Monticciolo, D. (1990). Medical audit of a rapid-throughput mammography screening practice: methodology and results of 27,114 examinations. ***Radiology***, 175, 323-327.

17. Linver, M., Paster, S., Rosenberg, R., Key, C., Stidley, C., & King, W. (1992). Improvement in mammography interpretation skills in a community radiology practice after dedicated teaching courses: 2-year medical audit of 38,633 cases. ***Radiology,*** 184, 39-43.

18. Elmore, G. J., Miglioretti, L. D., Reisch, M. L., Barton, B. M., Kreuter, W., Christiansen, L. C., & Fletcher, W. S. (2002). Screening mammograms by community radiologists: Variability in false-positive rates. ***Journal of National Cancer Institute***, 94, 1373-1380.

19. Thurfjell, E., Lernevall, K., & Taube, A. (1994). Bene¯t of independent double reading in a population-based mammography screening program. ***Radiology***, 191, 241-244.

20. Schmidt, R., Wolverton, D., & Vyborny, C. (1995). ***Computer-aided diagnosis (CAD) in mammography.*** A Categorical Course in Breast Imaging, 199-208.

21. Giger, L. M. 91996, June). Current issues in mammography. In ***Proceedings of the 3rd International Workshop on Digital Mammography*** *(*pp. 53-59). Chicago, IL.

22. Vyborny, C. & Giger, M. (1994). Computer vision and artificial intelligence in mammography. ***American Journal of Roentgenology,*** 162(3), 699-708.

*****

CLOUD COMPUTING SERVICES: AN OVERVIEW

Yougal Chandra Joshi[17] Suman Pathak[18]

ABSTRACT

This paper gives the insight of the cloud computing services from various angle of the published journal in this regard. Cloud computing simply mean to using a third-party network of remote servers hosted on the internet to store and manage data, rather than locally. On the other way round, **cloud services** provide the facility that makes any storage device such as hard disk in the cloud or on the internet. Such services have become popular because they are affordable, convenient and provide ample storage space. However, the important issue of such services is their accessibility, which provides access of documents, photos, videos and any other saved files from any device with internet access. With cloud, services one is can connect at home, work or on the go via a laptop, desktop, Smartphone or other handheld device. World's largest technical companies such as Apple, Amazon and Google have developed cloud services largely. These technical blue-chip companies, along with some notable runner-up, provide several storage tier plans customized for both consumers and businesses.

KEYWORDS

Cloud Computing, Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS) etc.

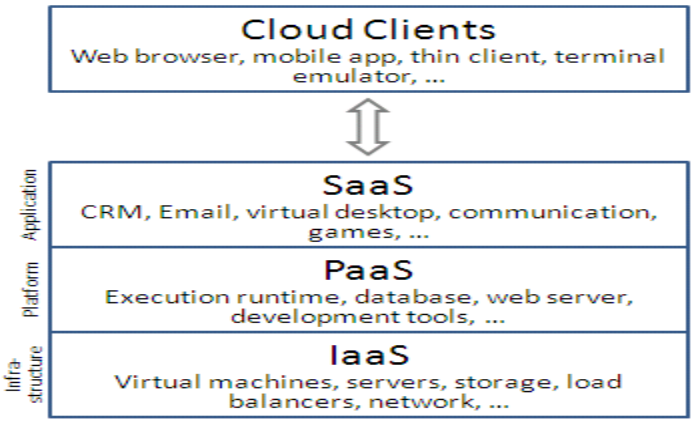
INTRODUCTION

Cloud computing describes highly scalable computing resources provided as an external service via the internet on a pay on demand basis. The cloud is simply a metaphor for the internet, based on the symbol used to represent the worldwide network in computer network diagrams. Economically, the main appeal of cloud computing is that customers only use what they need, and only pay for what they actually use. Resources are available to be accessed from the cloud at any time, and from any location via the internet. There is no need to worry about how things are being maintained behind the scenes. One's simply purchase the IT service as require, as would any other utility. Because of this, cloud computing has also been called utility computing, or 'IT on demand'. In fact, back in 1961, computing pioneer John McCarthy predicted "computation may someday be organized as a public utility" and went on to speculate how this might occur. Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services. The services themselves have long been referred to as Software as a Service (SaaS). The datacenter hardware and software is what we will call a Cloud. When a Cloud is made available in a pay on demand manner to the public, we call it a Public Cloud; the service being sold is Utility Computing. We use the term Private Cloud to refer to internal datacenters of a business or other organization, not made available to the public. Thus, Cloud Computing is the sum of SaaS and Utility Computing, but does not include Private Clouds. People can be users or providers of SaaS, or users or providers of Utility Computing. We focus on SaaS Providers (Cloud Users) and Cloud Providers, which have received less attention than SaaS Users [1].

SERVICE MODELS

Cloud computing providers offer their services according to three fundamental models: Infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) where IaaS is the most basic and each higher model abstracts from the details of the lower models [2].

Figure-1



Sources: Authors Compilation

[17]Information Scientist, Kumaun University, Uttarakhand, India, ycjoshi@yahoo.com
[18]Lecturer, Department of Management, Amrapali Institute of Management and Computer Application, Uttarakhand, India, sumanpathakjoshi@gmail.com

INFRASTRUCTURE AS A SERVICE (IAAS)

Infrastructure as a Service is a provision model in which an organization outsources the equipment used to support operations, including storage, hardware, servers and networking components. The service provider owns the equipment and is responsible for housing, running and maintaining it. The client typically pays on a per-use basis [3].

PLATFORM AS A SERVICE (PAAS)

Platform as a Service (PaaS) is a way to rent hardware, operating systems, storage and network capacity over the Internet. The service delivery model allows the customer to rent virtualized servers and associated services for running existing applications or developing and testing new ones. Platform as a Service (PaaS) is an outgrowth of Software as a Service (SaaS), a software distribution model in which hosted software applications are made available to customers over the Internet. PaaS has several advantages for developers. With PaaS, operating system features can be changed and upgraded frequently. Geographically distributed development teams can work together on software development projects. Services can be obtained from diverse sources that cross international boundaries. Initial and ongoing costs can be reduced by the use of infrastructure services from a single vendor rather than maintaining multiple hardware facilities that often perform duplicate functions or suffer from incompatibility problems. Overall expenses can also be minimized by unification of programming development efforts. On the downside, PaaS involves some risk of "lock-in" if offerings require proprietary service interfaces or development languages. Another potential pitfall is that the flexibility of offerings may not meet the needs of some users whose requirements rapidly evolve [4].

SOFTWARE AS A SERVICE (SAAS)

Software as a Service (SaaS) is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet. SaaS is becoming an increasingly prevalent delivery model as underlying technologies that support Web services and service-oriented architecture (SOA) mature and new developmental approaches, such as Ajax, become popular. Meanwhile, broadband service has become increasingly available to support user access from more areas around the world [5].

CLOUD CLIENTS

Users access cloud computing using networked client devices, such as desktop computers, laptops, tablets and smart phones. Some of these devices - *cloud clients* - rely on cloud computing for all or a majority of their applications to be essentially useless without it. Examples are thin clients and the browser-based Chrome book. Many cloud applications do not require specific software on the client and instead use a web browser to interact with the cloud application. With Ajax and HTML5, these Web user interfaces can achieve a similar or even better look and feel as native applications [6].

CONCLUSIONS

The future of cloud computing is moving toward more ubiquity, as greater demands from customers and greater capabilities from providers unfold. It is the beginning of a new Internet based service economy, the Internet centric, Web based, on demand, Cloud applications and computing economy. Cloud Computing is a general term that hinges over a variety of services from Infrastructure as a Service at the base, through Platform as a Service as a development tool and through to Software as a Service replacing on-premise applications.

REFERENCES

1. Ian, Foster, Yong, Zhao, Ioan, Raicu, & Shiyong, Lu. ***Cloud Computing and Grid Computing 360-Degree Compared,*** 1-10.
2. Retrieved on August16, 2012 from http://en.wikipedia.org/wiki/Infrastructure_as_a_service#Service_models
3. Retrieved from http://searchcloudcomputing.techtarget.com/definition/Infrastructure-as-a-Service-IaaS
4. Retrieved from http://searchcloudcomputing.techtarget.com/definition/Platform-as-a-Service-PaaS
5. Retrieved from http://searchcloudcomputing.techtarget.com/definition/Software-as-a-Service-SaaS
6. Ibid, 2.

*****

WOMEN AND INFORMATION COMMUNICATION TECHNOLOGY

Shweta Chaturvedi[19]

ABSTRACT

'It has been observed that many of Indian women are not able to connect with ICT because of financial dependency, lack of English language.lac of technical knowledge and lack of resources like light, teaching staff and a very important factor that they don't know the importance of ICT too. Even every big telecom company is focusing on urban areas that are why people of rural areas don't have access to ICT resources.'

The whole research paper is divided into three parts. In first part, relation between women and ICT is detailed. In second part, the basic problem, which is faced by the women whether they live in urban area or rural areas, will be also explained, and in third part of this research paper solutions of these problems will be given with the help of some suggestions.

KEYWORDS

ICT, Women, International , Rural, Urban etc.

INTRODUCTION

Information & Communication Technology (ICT) is an umbrella term that includes any communication device or application, like as radio, television, mobile phones, computer and network hardware and software, satellite system and so on, as well as the various services and application associated with them, such as Video Conferencing and distance learning. So, ICT is concerned with the storage, retrieval, manipulation, transmission of digital data. Information & Communication Technology drives a competitive marketplace. It still provides our best basis of co-operation, ICT drive modern culture, ICT offer new opportunities for diplomacy, for social development, and for better public administration through advance in e-governance on the internet.

There has been a rapid growth in the ICT sector since the late 1980s and the use of ICT has expanded since 1990s. According to the World bank Teledensity in India had reached 3.8% by 2001, and the number of internet accounts is growing a rate of 50% per annum. ITs and BPO sector grew at 59% but there is a strong digital divide in the society. Sam Pitroda described that the ICT revolution in India is about to enter into its second phase that would be based on broadband, data, applications, process and new infrastructure. This would change the nature of education, health, government services, and public delivery system available in India. Sam added that in today's time, nothing is possible without IT and ICT. IT offers unique possibilities to make generation change.

OBJECTIVES OF STUDY

The objective of this research paper is to find out the real impact of information communication technology on women at national and international level.

REVIEW OF LITERATURE

It has been observed that many of Indian women are not able to connect with ICT because of financial dependency, lack of English language.lac of technical knowledge and lack of resources like light, teaching staff and a very important factor that they do not know the importance of ICT too. Even every big telecom company is focusing on urban areas that are why people of rural areas do not have access to ICT resources. In addition, why they can't access it? The reason is that because the basic need to access is not there, they lack of infrastructure and facilities, there are poll wires but there is no electricity, literacy is also the main problem, majority of the population does not know how to access ICT, and if they knows how to access it they don't know what to access, even in developing countries people who access ICT they watch pornography.

Since India has been using ICT for development for more than two decades, there are many good practices for the use of ICTs for women's empowerment. ICT has played an important role in changing the concept of work and workplace, new area of employment such as tele-working, i.e. working from distance, are becoming feasible with new technology .Female are always discriminated in the field of economic, education, health and social services worldwide. On the other hand the range of women's economic activities in developing countries is improving .It includes formal, informal and self-employment in farming, trading and craft production. There are the possibilities for ICT to improve women's economic activities in the field of trade, governance, education, health, crafts employment in formal as well as informal sector.

CASE STUDIES

ICT's bring many opportunities to women in the work situation and small business. In 2005, in Tamil Nadu an e-commerce website has been designed to sell products made by rural women's co-operative and NGO's. The Dhan foundation and Swayam

[19] Assistant Lecturer, Lingaya Lalita Devi Institute of Management and Sciences, New Delhi, India, shweta_chaturvedi11@yahoo.com

Krishi sangam are using ICT's such as handheld devices and smart cards to improve microfinance projects to empower poor women; the self-employed women's association has several ICT projects for women.

You can earn up to 50000 per month and that too by sitting at home. People can purchase stuff from market just by dialing a number, they can see their relative sitting in abroad, they can transfer the money anywhere in the world just by a click, earlier these were sounding strange but in present days it is common with the help of Information and communication technology, ICT consist of IT as well as telecommunication, broadcast media, all types of audio and video processing and transmission and network based control and monitoring functions. Countries have recognized ICT as an effective tool in catalyzing the economic activity in efficient Governance and in developing human resources.

IT with communication technologies has brought about drastic changes in the way people communicate. The emergence of IT with communication makes the lives of people much better and more comfortable in several ways. Information is the key to democracy; with the advent of ICT, it has become possible for the common person to access global information. The realm of electronic communication encompasses telecommunication, broadcasting, and information technology. Information is now broadly defined to embrace voice in telephony, text in fax and newspaper, images in video and television broadcasting and data in computers. All information can be digitized, transported, stored, retrieved, modified and distributed.

According to Blanca, ICT have created new jobs in the field of banking, publishing, insurance, printing, and that too especially for women. Maximum media house comprise of women, even in media women are holding key positions. Anurradha Prasad the editor in chief of BAG films, Alka Saxena editor of ZEE news, Barkha Dutt, Group Editor with NDTV, Nalini Singh, Chief Editor of TV LIVE, Mrinal Pande, Chief Editor of HINDI DAILY, HINDUSTAN are few of the names holding a high position in electronic media.ICT changing life, living standard, thought, ideas, status of women in the society. In modern India women have taken charges at high offices in the country, As of 2011 , the president of India, the speaker of the Lok Sabha and the leader of the opposition in Lok Sabha are all women but this is just one side of the coin because people of urban area are getting benefit.

The literacy rate of women has improved but not at the satisfactory level, they are taking their educations equally as men but at some extent, there are majority of the women who are not getting proper education, they are working in all the fields like political, literature, music and art, banking, insurance they are everywhere and some of them are sitting at the higher position. They participate in social works, they fight for others, they save others life. They are equally working as men, and they are technically advanced. The female population of the rural areas are getting computer classes, and coaching. Even a woman sitting in a village knows how to use a mobile and literally enjoys it but only by using mobile, we cannot say that they are technically advanced. There is a lot to do for them to make them technically advanced.

Earlier women were not treated equally as men, they were treated like they are the property of men, they were not suppose to talk to any men accept her family members, they were not suppose to step out of their homes, and they were not given any responsibilities. However, with the time they changed the mind of the society, they started proving themselves, they participated in the freedom struggle, when men folk were in prison then a remarkable thing happened, women came forward and took charge of the struggle. In the women of modern era is very different from past. Women are showing their presence in every field, like sports, science, literature, art etc but that is in minority.

Women of rural areas are far behind than urban areas. They are not more educated, they do not have will power to fight for social justice, they do not have financial support, they do not have freedom to speech, and they do not have knowledge of government plan for women empowerment. However, the women of urban areas are much more aware for her rights, freedom, social justice, health, wealth etc. they also fight for others. They have reach and access to the ICT. They know each and every government plan like as Ladli for girls, women empowerment plan, right to information, right to protect from domestic violence etc. they are well aware of all these because of ICT.

WOMEN OF WESTERN WORLD

Even the comparison with the women of western world, Indian women are not in picture. Their growth rate, education, knowledge of computer, English speaking skills making them more empowered than Indian women. Most of them use to travel all over the world, they are not financially dependent, and they use more ICT for their work.

In the Arabian countries, women are not much aware of ICT. Social and religious factors are responsible for lack of ICT in the Muslim countries. They do not have positive thinking for women; they are rivalry of women empowerment. Women of Muslim countries are spending their life under the burkas. Majority of the women are not equally educated as men, men are still dominating in the economic development, and men have full responsibilities of the family .In some nations women are prohibited from travelling alone.

CONCLUSIONS

Women can come at the same level and they should come at the same level, they should get the same benefit because every woman is special and important. Women can be equally treated and can get the same benefit by some solutions. They should get the basic infrastructure and facilities; if the government provides them basic needs like food, shelter, clothing, health and education then definitely women can be saved and can be educated as education is very important to change the mentality and the thought process.

If the technology is not expensive then every common man can access it very easily, illiterate women should get chance to talk to educated women, it will help them to improve, it will remove their hesitation by which they can be freely expressive, and when these things are transformatted. It is important because women are the creature of the world hence they should be taken care, they are special.

REFERENCES

1. Mishra, C. R. ***Teaching of information technology***. ISBN No. 81-7648-890-9, APH Publication.
2. Rao, Pramod, & Gore, J. Medha. ***Women Entrepreneurs***. ISBN No. 978-81-314-1907-6, The ICFAI University press.
3. Varma, Anil. ***Information and communication Technology in Education***. ISBN No. 978-81-314-1629-7, The ICFAI University press.
4. Saxena, Prakash Rajiv. ***Government of the future in the ICT Era.*** ISBN No. 81-314-0109-X, The ICFAI University Press.
5. Sharma, M. B. ***Women and Education.*** ISBN No. 81-7169-308-3, Commonwealth Publishers.
6. Narasaiah, L. M. ***Information Technology and Development***. ISBN No. 81-7141-760-4, Discovery Publishing House.
7. Mukherjee, Debashree. ***Women Education and Empowerment.*** ISBN No. 978-81-314-1568-9, The ICFAI University Press.
8. Multani, Kaur Sukhvinder. ***ICT in Rural Development.*** ISBN No. 81-314-0833-7, The ICFAI University Press.
9. Nag, Paramita. ***Violence against Women***. ISBN No. 978-81-314-2164-2, The ICFAI University Press.

*****

ONTOLOGY: AN ANALYTICAL STUDY

Ashish Kumar[20]

ABSTRACT

In the age of information explosion, improved knowledge management is crucial; and to get successful. Information communication technology is playing a pivotal role. The potential of information technology in knowledge management is enormous, many tools and techniques are being developed to achieve this goal. Ontology is one of the tools developed for knowledge management. It is basic description of things in the world. In the paper, comprehensive introduction of ontology is provided, its use, and potential is discussed to make Information professionals alert about Ontology and its applications.

KEYWORDS

Information Retrieval, Knowledge Management, Database Modeling, Ontology, Semantic Web, Ontology, Digital Library, Artificial Intelligence etc.

INTRODUCTION

Information is the backbone of development in this era, many technologies has developed for storage, manipulation and retrieval of information. Development in these technologies has opened a new vista of retrieval and use of information; due to advancement in this field, many new terms have been developed to represent concepts. Ontology is one of the terms, which is being used in subjects like Computer science, Semantic Web, Knowledge Engineering, Knowledge Management, Artificial Intelligence and Semantic Interconnection. While different subject professionals to refer their own subject's expression are using the term "Ontology", the term is also used in the library and information science. Due to different expressions by different professionals, sometimes, library science professionals become confused about the context of the term.

The paper is an attempt to trace the history of the term "Ontology", the relationship of ontology with other subjects like Computer science, Internet and Library and Information science through review of the literature. In the paper an overview of Ontology, in terms of its definition and historical background is defined. In next section, Ontology in Computer science and Internet is discussed and finally, the significance of ontology in context to Library and Information science is examined in the paper.

ONTOLOGY: AN ANALYSIS

The term Ontology, literally means "the branch of metaphysics that studies the nature of existence or being as such." It takes its derivation from Latin terms '*onto'* – being, and *ology* - 'the study of'. (Random-House, 2000) Ontology, as a discipline of thought, is a principle of metaphysics which includes cosmology being the study of the physical universe, *cosmo*s –universe and *ology* –the study of, and also is intimately connected to epistemology which is the study of knowledge, *epistem* –knowledge and *ology* –the study of.[1] (Random-House, 2000).

Ontology, as a branch of philosophy, is the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality. Philosophers often use 'Ontology' as a synonym of 'metaphysics', a term used by Aristotle to refer 'first philosophy'.

Sometimes 'ontology' is used in a broader sense, to refer to the study of what *might* exist; 'metaphysics' is then used for the study of which of the various alternative possible ontologies is in fact true of reality. (Ingarden 1964) [2]. Two philosophers coined the term 'ontology' or ontologia in 1613, independently, Rudolf Göckel (Goclenius), in his *Lexicon philosophicum* and Jacob Lorhard (Lorhardus), in his *Theatrum philosophicum*. Its first occurrence in English as recorded by the OED appears in Bailey's dictionary of 1721, which defines ontology as 'an Account of being in the Abstract'. In the philosophical tradition, ontology is related to what exists a priori to perception, knowledge, or language. It is sometimes further divided in reality-based and epistemological ontology; the latter describes human conceptualizations of reality (Smith 1999) [5]

ONTOLOGY AND COMPUTER SCIENCE

The notion of ontology was borrowed from philosophy and adopted in Computer Science under multiple definitions, where, probably, the first credible and the most commonly quoted one is "an explicit specification of a conceptualization." [3]

Early Artificial Intelligence (AI) researchers, who recognized the applicability of the work from mathematical logic [8] and argued that AI researchers could create new Ontologies as computational models that enable certain kinds of automated reasoning, had adopted the term. In the 1980s, the AI community came to use the term ontology to refer to both a theory of a modeled world [7] and a component of knowledge systems. Some researchers, drawing inspiration from philosophical Ontologies, viewed computational ontology as a kind of applied philosophy [9]. In computer and information science, ontology is a technical term denoting an artifact that is *designed* for a purpose, which is to enable the modeling of knowledge about *some* domain, real or

[20]Assistant Librarian, Kumaun University, D.S.B. Campus, Uttaranchal, India, ashish3k@gmail.com

imagined. [6] In the early 1990s, an effort to create interoperability standards identified a technology stack that called out the ontology layer as a standard component of knowledge systems. [10]

A widely cited web page and paper is credited with a deliberate definition of ontology. According to Gurbur, ontology is a technical term in computer science. [3] He defines ontology as an "explicit specification of a conceptualization," which is, in turn, "the objects, concepts, and other entities that are presumed to exist in some area of interest and the relationships that hold among them."[11]. While the terms specification and conceptualization have caused much debate, the essential points of this definition of ontology are:

- Ontology defines (specifies) the concepts, relationships, and other distinctions that are relevant for modeling a domain.
- The specification takes the form of the definitions of representational vocabulary (classes, relations, and so forth), which provide meanings for the vocabulary and formal constraints on its coherent use.[6]

Gurbur further explained the term conceptualization, as "A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose. Every knowledge base, knowledge-based system, or knowledge-level agent is committed to some conceptualization, explicitly or implicitly." [11]

An assortment of techniques for representing and managing codified knowledge has emerged from a number of areas in computer science, notably artificial intelligence, databases, software engineering, and information systems. This movement towards knowledge orientation has not been organized as a coherent movement or even viewed as such, as it has come about for a variety of reasons. From a practical standpoint, the growing complexity of application domains, of software development, and the increasing intertwining of machine and human processes have all contributed to such needs and the development of techniques to address them. (Bubenko, 1980; Newell, 1982; Ullman, 1988). [34] The key role of Ontologies with respect to database systems is to specify a data modeling representation at a level of abstraction above specific database designs (logical or physical), so that data can be exported, translated, queried, and unified across independently developed systems and services. Successful applications to data include database interoperability, cross database search, and the integration of web services. [6]

ONTOLOGY AND INTERNET

World Wide Web (WWW) grew mainly around HTML. It is a huge library of interlinked web pages that are transferred by computers and presented to people. It has grown from hypertext systems, which provides a standard for structuring web pages that browsers can translate to retrieve documents. Due to the limitation of HTML, another language XML has developed. XML is a defined way to provide a serialized syntax for tree structures; it is an important first step toward building a Semantic Web, where application programs have direct access to data semantics.

The **Semantic Web** is a "web of data" that enables machines to understand the semantics, or meaning, of information on the World Wide Web.[33] In Semantic web, meaning is expressed by RDF (a component of web), which encodes it in sets of triples; each triple are like the subject, verb and object of an elementary sentence. These triples can be written using XML tags. In RDF, a document makes assertions of particular things (people, Web pages or whatever), have properties (such as "is a sister of," "is the author of") with certain values (another person, another Web page).This structure turns out to be a natural way to describe the vast majority of the data processed by machines. A Universal Resource Identifier (URI) identifies each subject and object, just as used in a link on a Web page. The verbs are also identified by URIs, which enables anyone to define a new concept, a new verb, just by defining a URI for it somewhere on the Web. The URIs ensures that concepts are not just words in a document but are tied to a unique definition that everyone can find on the Web. Different databases may use different identifiers for what is, in fact, the same concept. A program that wants to compare or combine information across the two databases has to know that these two terms are being used to mean the same thing. Ideally, the program must have a way to discover such common meanings for whatever databases it encounters. [35]

A solution to this problem is provided by the third basic component of the Semantic Web, Ontologies. Ontology is a document or file that formally defines the relations among terms. The most typical kind of ontology for the Web has taxonomy and a set of inference rules. The taxonomy defines classes of objects and relations among them. Classes, subclasses and relations among entities are a very powerful tool for Web use. We can express a large number of relations among entities by assigning properties to classes and allowing subclasses to inherit such properties. The computer does not truly "understand" any of the information and its relation, but it can manipulate the terms much more effectively in ways that are useful and meaningful to the human user. Ontologies can enhance the functioning of the Web in many ways. They can be used in a simple fashion to improve the accuracy of Web searches—the search program can look for only those pages that refer to a precise concept instead of all the ones using ambiguous keywords. More advanced applications will use Ontologies to relate the information on a page to the associated knowledge structures and inference rules. [35]

Use of ontology has become common on the WWW due to sheer volume and variety of information items. The Ontologies on the Web currently range from large taxonomies, categorizing websites to grouping of products according to their nature to facilitate search. Support of Ontologies for both browsing and searching the digital information is of tremendous assistance. For specific as well as for general purpose, Ontologies have been developed to help the subject experts and others to share and annotate information in the corresponding field. Aim is to share and communicate knowledge, between both people and software agents. [3] Sharing common understanding of the structure of information among people or software agents is one of the more common goals in developing Ontologies. (Musen 1992)[13]

Ontology, with its inference mechanism, is the key technology of semantic web. With the development of semantic web, ontology is widely used in many domains. Ontology provides a brand new model of knowledge management, which solves the problems of knowledge organization, knowledge retrieval and so on [27].

Users to find their actual and potential search requirements step by step which depends on two methods; one is using edit distance based method to search word in ontology base and the other is using concept space-based method to help user to extend keywords could use Ontologies. The benefit of the second method is that it utilizes the relations between words built by field ontology to find word's actual meanings in its context. [28] According to Saravanan, Ravindran and Raman, ontological frameworks can be used to improve retrieval performance. [30]

Ontologies specifically play an important role in supporting knowledge-based applications in semantic web. [24] They are the backbone of a semantic web, a semantic version of the World Wide Web. [25] The success of the semantic Web depends strongly on the proliferation of Ontologies. [26]

ONTOLOGY AND INFORMATION SCIENCE

Libraries are the learning institutions from the beginning of their existence. Without libraries, learning would have been impossible. They acquire, organize, retrieve and disseminate knowledge through their services. Traditionally, the services and tools such as Classification, Cataloguing and Indexing are developed for paper-based documents. Documents have traditionally been paper-based and their contents can be read, interpreted and implemented by their users. Advancement in Information and Communication Technology has changed the information scenario resulting different Information objects such as e-journal, e-book, which could range from a single webpage to a pixel-based photograph to a digital piece of music.

ICT has also influenced the concept of libraries and now we have a different format of libraries; among these, we can find hybrid library, digital library and virtual library. An informal definition of a digital library is "a managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network." (Arms, 2002) [36] The hybrid library provides electronic information sources too but it is paper-based information. The hybrid should be considered as a model by itself not as a transitional phase from a conventional library to a digital one (Brophy, 2001).[37] And finally the virtual library has been defined as the concept of a "remote access to the contents and services of libraries and other information resources, combining an on-site collection of current and heavily used materials in both print and electronic form, with an electronic network which provides access to, and delivery from, external worldwide library and commercial information and knowledge sources." (Gapen, 1993) [38].

The contents of documents of these libraries are also individually meaningful. However, to understand them in broader and deeper context one requires other connected documents, or some kind of mechanism to connect the documents. Representations of learning resources must support the finest-grained level of granularity required by the core technologies as suggested in [18], in addition to application and support technologies. Learning objects [17] within a learning resource need to be encoded in a way that they can be recognized, searched, referenced, and activated at different granularities. Other researchers have been addressing related technological solutions, such as dynamic metadata and automated component descriptions. [19, 20]

Despite several advancements in the tools and services of libraries, the basic approach of retrieval remains very much similar; the bibliographic description approach. Generally, all information systems use two distinct schemas for encoding information embodied in a document. One represents the subject contents of a document using a classification scheme, a thesaurus or a subject-heading list. The other provides descriptive information about a document like title, author, publisher, etc. These are handled by metadata schema such as cataloguing codes, MARC and other bibliographic description formats and more recently Dublin Core [31]. Dublin Core (DC) [21] provides basic factual description, which is used most commonly in creating collection or resource level metadata.

While data and information are captured and represented in various digital formats, and rapidly proliferating, the techniques for accessing data and information are rudimentary and imprecise, mostly based on simple keyword indexes and relational queries. In the current context of explosive availability of data, there is a need for a knowledge discovery approach, based on both top-down knowledge creation (e.g., Ontologies, subject headings, user modeling) and bottom-up automated knowledge extraction (e.g., data mining, text mining, web mining), promises to help transfer digital library from an institution of data and information to an institution of knowledge. (Chen, 2003) [39]

B.C. Vickery first drew attention to the concept of ontology for organizing knowledge in the wake of its increasing complexity. [32] Ontology, in the field of information management, defines a common vocabulary for users who need to share information in a domain. [3] Ontology is a representation vocabulary, often specialized to some domain or subject matter. More precisely, it is the conceptualizations that the terms in the vocabulary are intended to capture.[12] The distinguishing feature is that it includes machine-interpretable definitions of basic concepts and relations among them.[3]

At development time, an ontology can be used in the conceptual modeling phase of IS, representing the knowledge of a given domain and supporting the creation of IS components, such as information resources, applications programs, and user interfaces.

Ontology allows computer to process information resources based on their contents. Ontology is seen as describing subject knowledge matter using concepts, relations, functions and assumptions in addition to a taxonomy. A typical ontology for the Web will have a taxonomy and set of inference rules. In this context, taxonomy is treated as 'simple ontology' and subject heading lists

and thesauri as 'lightweight Ontologies' in some quarters. Classification provides say a container to put the information items, i.e., where the data are, but ontology provides further information about their relationships, i.e., what the data are. [22] Ontologies ensure efficient retrieval by enabling inferences based on domain knowledge, which is gathered during the construction of knowledge base [23].

Building Ontologies for Information Systems is not an easy task, and requires a great set of skills of Ontology Engineering. There has been a proliferation of conceptual modeling methods that use ontology as an artifact in Information Systems Analysis and Design (ISAD). [40, 41, 42, 43]

DISCUSSIONS

Ontology is being used in different field as the analysis in the field of computer science, Semantic web and library and information science is already provided, other disciplines also are developing standardized Ontologies that domain experts can use to share and annotate information in their fields. Medicine, for example, has produced large, standardized, structured vocabularies such as SNOMED (Price and Spackman 2000) [14] an architecture for an ontology-driven geographic information system (ODGIS). The system uses an object-oriented mapping of Ontologies. [16]

Handschuh et al. provide a framework, known as S-CREAM, which allows for creation of metadata and is trainable for a specific domain. Ont-O-Mat is the reference implementation of the S-CREAM framework. It provides a plugin interface for extensions for further advancements, e.g., collaborative metadata creation or integrated ontology editing and evolution. [47]

Separating the domain knowledge from the operational knowledge is another common use of Ontologies. One can describe a task of configuring a product from its components according to a required specification and implement a program that does this configuration independently of the products and components themselves (McGuinness and Wright 1998) [15]

Ontologies are also used in system designed. OntoSeek is a system designed for content-based information retrieval. [45] A method of extracting information from unstructured documents based on application ontology is presented. Embley et al. [46] Alani et al. propose the Artequakt that automatically extracts knowledge about artists from the web based on ontology. It can generate biographies that tailor to a user's interests and requirements. [48] Lee et al. propose an ontology-based fuzzy event extraction agent for Chinese news summarization. The summarization agent can generate a sentence set for each Chinese news, but the results need to be improved to suit for various Chinese news websites. [49] Halteren develops a new technique to find out several feature sets through the set of style makers and extracts summarization sentences from the document, according to the found feature sets for users. [50]

Vargas-Vera et al. present an annotation tool, called MnM, which provides both automated and semi-automated support for annotating web pages with semantic contents. MnM integrates a web browser with an ontology editor and provides open application programming interfaces (APIs) to link to ontology servers and for integrating information extraction tools. The goal of text summarization is to take the abstract from the extracted content and present the most important message for the user in a condensed form. [31] Navigli et al. propose the OntoLearn with ontology learning capability to extract relevant domain terms from a corpus of text. [44] Lam et al. propose the Financial Information Digest System (FIDS) to summarize online financial news articles automatically. The FIDS can integrate the information from different articles by conducting automatic content-based classification and information item extraction. [29]

CONCLUSIONS

From the literature, it has been found that Ontology has a potential to change the way we see and use the information, it does have the capability to change the WWW specifically and Information retrievals system in general. Despite its promises, many researchers feel that one of the major challenges in the area of ontology design is creating the contents of Ontologies; Ontology evaluation also remains an important problem in the area of ontology-supported computing and semantic web. There is no single approach to ontology evaluation and it more or less depends upon the purpose of which it is being developed. There is a need of uniform evaluation method to judge Ontologies.

REFERENCES

1. ***Random House Webster's College Dictionary***. (2000). New York: Random House.

2. III. Charles, Thomas. (1964). Extracts from a masterly four-volume work in realist ontology entitled ***The Problem of the Existence of the World*** (Helen R. Michejda, Trans.). In Ingarden Roman, ***Time and Modes of Being*****,** Springfield, Ill.

3. Gruber, R. T. (1993). A translation approach to portable ontology specifications**.** ***Knowl. Acquis***. 5(2), 199-220.

4. Floridi, L. (Ed.). (2003). ***Blackwell Guide to the Philosophy of Computing and Information*****,** Ontology. Oxford: Blackwell, 155–166.

5. Smith, B. (1999). An introduction to ontology. In ***The Ontology of Fields, Report of a Specialist Meeting held under the auspices of the Varenius Project***, (Edited by D. Peuquet, B. Smith and B. Brogaard), NCGIA, Bar Harbour, Maine.

6. Ling, Liu, & Tamer, Ozsu M. (2008). ***Ontology to appear in the Encyclopedia of Database Systems*** (Ed.). Springer-Verlag.

7. Hayes, P. J. (1985). The Second Naive Physics Manifesto. In Hobbs and Moore (eds.). ***Formal Theories of the Common-Sense World***. Norwood: Ablex.

8. McCarthy, J. Circumscription. (1980). A Form of Non-Monotonic Reasoning. ***Artificial Intelligence,*** 5(13), 27-39.

9. Sowa, J. F. (1984). ***Conceptual Structures. Information Processing in Mind and Machine.*** Reading, MA: Addison Wesley.

10. Neches, R., Fikes, R. E., Finin, T., Gruber, T. R., Patil, R., Senator, T., & Swartout, W. R. (1991). Enabling technology for knowledge sharing. ***AI Magazine,*** 12(3), 16-36.

11. Gruber, T. R. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. ***International Journal Human-Computer Studies*****,** 43(5-6), 907-928.

12. Chandrasekaran, B., & Josephson, John R. (1999, January-February). ***What Are Ontologies, and Why Do We Need Them?*** University of Amsterdam (IEEE Intelligent Systems).

13. Musen, M. A. (1992). Dimensions of knowledge sharing and reuse. ***Computers and Biomedical Research,*** 25**,** 435-467.

14. Price, C., & Spackman, K. (2000). SNOMED clinical terms. ***BJHC&IM-British Journal of Healthcare Computing & Information Management,*** **17**(3), 27-31.

15. McGuinness, D. L., & Wright, J. (1998). Conceptual Modeling for Configuration: A Description Logic-based Approach. ***Artificial Intelligence for Engineering Design, Analysis, and Manufacturing-Special issue on Configuration*****.**

16. Frederico, T. Fonseca, & Max, J. Egenhofer. (1999, November). Ontology-Driven Geographic Information Systems. In ***7th ACM Symposium on Advances in Geographic Information Systems Kansas City***, MO C. Bauzer Medeiros (ed.).

17. (2012, November 18). ***IMS Learning Resource Metadata***. Retrieved from http://www.imsproject.org/metadata/

18. Laleuf, J. R., & Spalter, A. M. (2001, June 24-28). A Component Repository for Learning Objects: A Progress Report. ***JCDL***, Roanoke, VA.

19. Damiani, E., Fugini, M. G., & Bellettini, C. (1999). A Hierarchy-Aware Approach to Faceted Classification of Object-Oriented Components. ***ACM Transactions on Software Engineering and Methodology,*** 8(3), 215-262.

20. Saddik, A. E., Fischer, S., & Steinmetz, R. (2001). Reusability and Adaptability of Interactive Resources in Web-Based Educational Systems. ***ACM Journal of Educational Resources in Computing*** 1(1).

21. (2012, November 21). Dublin Core Metadata Initiative Retrieved from http://dublincore.org

22. Gilchrist, A. (2003). Thesauri, taxonomies and Ontologies: An etymological note. ***Journal of Documentation*****, 59**(1), 7-18.

23. Saravanan, Ravindran & Raman. (2009). Improving legal information retrieval using an ontological framework. ***Arti. Intell. Law***, **17**(2), 101-24.

24. Jiang, & Tan. CRCTOL: A semantic-based domain ontology learning system. ***J. Amer. Soc. Inf. Sci. Technol.***, **61**(1), 150-168.

25. De Nicola, Missikoff & Navigli. (2009). A software engineering approach to ontology building. ***Information Systems,*** 34(2), 258-275.

26. Lin, & Hanqing. (2009). Research on semi-automatic construction of domain ontology based on statistical NLP technique. ***J. China Soc. Sci. Tech. Inf,*** **28**(2), 201-207.

27. Hui (*et. Al.).*(2009). Research on digital library based ontology evolution and knowledge management (I) ontology molecule theory. ***J China Soc. Sci Tech. Inf.*** **28**(3), 323-330.

28. Mingbao, Jing, & Xiuli. (2010). Research on field ontology applications in information retrieval. ***J. China Soc. Sci. Tech. Inf.,*** **29**(2), 215-222.

29. Lam W. & Ho, S. K. (2001, November). FIDS: An intelligent financial web news articles digest system. ***IEEE Trans. Syst.,*** Man, Cybern. A, 319(6), 753–762.

30. Saravanan, Ravindran, & Raman. (2009). Improving legal information retrieval using an ontological framework. ***Arti. Intell. Law,*** **17**(2), 101-124.

31. Vargas-Vera, M., Motta, E., Domingue, J., Lanzoni, M., Stutt, A. & Ciravegna, F. (2002). MnM: Ontology driven semi-automatic and automatic support for semantic markup. In ***Proc. 13th Int. Conf. Knowledge Eng. Manage.***, 379–391.

32. Vickery, B.C. (1993). Ontologies. ***J. Inf. Sci.,*** **23**(4), 276-286.

33. ***W3C Semantic Web Frequently Asked Questions.*** Retrieved on December 4, 2012 from http://www.w3.org/Help/#activity.

34. Bubenko, J. A. (1980). Information Modeling in the Context of System Development. In ***Proceedings IFIP Congress*** (pp. 395-411).

35. Berners-Lee, Tim, Hendler, James, & Lassila, Ora. ***The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities.*** Retrieved on December 2, 2012 from http://www.cs.umd.edu/~golbeck/LBSC690/SemanticWeb.html

36. Arms, W.Y. (2002). ***Digital Libraries.*** Cambridge, MA: The MIT Press.

37. Brophy, P. (2001). ***The Library in the Twenty-First Century***. London: Library Association Publishing.

38. Gapen, D. K. (1993). The virtual library: knowledge, society and the librarian. In Saunders, L. (Ed.), ***The Virtual Library Visions and Realities*** (1-14). Meckler Corporation, Westport, CT.

39. Chen, H. (2003). ***Towards building digital library as an institution of knowledge.*** NSF Post Digital Library Futures Workshop. Chatham, MA. Retrieved from www.sis.pitt.edu/%7Edlwkshop/paper_chen.html

40. Guarino, N. (1998). Formal Ontology in Information Systems. In ***Proceedings of the FOIS'98.*** Trento, Italy, Amsterdam, IOS Press.

41. Bubenko, J. A. (1979). On the Role of Understanding Models' in conceptual schema design. In ***Proceedings of the Fifth International Conference on Very Large Data Bases***, Rio De Janeiro, Brazil, IEEE.

42. Fonseca, F. (2007). The Double Role of Ontologies in Information Science Research. ***Journal of the American Society for Information Science and Technology***, 58(6), 786-793.

43. Wand, Y., & Weber, R. (1989). An Ontological Evaluation of Systems Analysis and Design Methods. In ***Proceedings of the Information Systems Concepts: An in-Depth Analysis*****,** Elsevier Science Publisher.

44. Navigli, R., & Velardi, P. (2003, January-February). Ontology learning and its application to automated terminology translation. ***IEEE Intell. Syst.,*** 18(1), 22–31.

45. Guarino, N., Masolo, C., & Vetere, G. (1999, May-June). Onto Seek: Content-based access to the web. ***IEEE Intell. Syst.,*** 14(3), 70-80.

46. Embley, W. D., Campbell, M. D., Smith, D. R., & Liddle, W. S. (1998). Ontology-based extraction and structuring of information from data-rich unstructured documents. In ***Proc. ACM Conf. Inf. Knowledge Manage.*** (pp. 52-59).

47. Handschuh, S., Staab, S., & Ciravegna, F. (2002). S-CREAM-semi-automatic Creation of metadata. In ***Proc. 13th Int. Conf. Knowledge Eng. Manage.,*** (pp. 358–372).

48. Alani, H., Kim, S., Millard, E. D., Weal, J. M., Hall, W., Lewis, H. P., & Shadbolt, R. N. (2003, January/February). Automatic ontology-based knowledge extraction from web documents. ***IEEE Intell. Syst.,*** 18(1), 14–21.

49. Lee, S. C., Chen, J. Y., & Jian, W. Z. (2003). Ontology-based fuzzy event extraction agent for Chinese e-news summarization. ***Expert Systems with Applications***, 25(3), 431-447.

50. Halteren, V. H. (2003, July-August). New feature sets for summarization by sentence extraction. ***IEEE Intell. Syst.,*** 18(4), 34-42.

*****

E-LEARNING AND M-LEARNING FRAMEWORK USING CLOUD BASED SERVICES

Prachi Dhaulakhandi[21] Parveen Sharma[22]

ABSTRACT

Software as a Service has been with us for some time like use of email services. This is a decades old concept. Educators to provide online contents have been using cloud-hosted services to deliver education for over a decade. This paper presents a new Model of Distance Learning System to offer e-learning and m-learning based modules in Cloud Computing Environment, by using High Performance cloud Computing (HPCC) Infrastructure, as well as some existing videoconferencing technologies enriched with mobile and wireless devices. This paper analyzed the difference that what a user should expect from the cloud: in terms of features, security, redundancy, scalability, automation and cost reduction.

KEYWORDS

m-Learning , e-Learning ,Cloud Services, SaaS , OLMS , Cloud Security etc.

INTRODUCTION

From the user standpoint, the concept of cloud computing can best be explained as a collection of server delivering resources that can be accessed remotely via the Internet in real-time. These servers are housed in a bunker like structure called a Data Center. In other words, your data, your software applications are not housed on your computer; they are on a service's cloud of web servers (often virtual servers) usually accessed by you via the Internet using a browser like Chrome or IE. You are renting the use of the software and storage space. The cloud is effectively a group of servers; more specifically– "virtual servers"– that simulate running multiple computers on a single piece of hardware. This is beneficial since it is possible to get more use out of the piece of hardware than if it was just doing the work of one. A simple explanation: If I have ten Dell servers, each at 10% utilization I will have ten physical pieces of equipment to maintain and upgrade or using the cloud, I can have only one server at 100% utilization. The term cloud leverages the fact that these virtual servers can be started up, shut down, upgraded, moved from physical machine to physical machine, etc all through software and in response to demand or other event. For example, you might want to have more web servers running during the day when traffic is high and fewer during the evening when traffic is low. Typically, cloud servers cost out per hour. This can be more economical than keeping all of your web servers running all of the time.

CLOUD COMPUTING FOR E-LEARNING AND M-LEARNING

Clouds are defined by the technology they provide: computation, software, data access, and storage services. A cloud can be defined as a place for users to create or store files, but has alternative meanings that, for example, explain how using a cloud can optimize processing power on the user end through its network. Services now deliver software such as Microsoft Office from the cloud. This means a computer user is renting the use of the software- usually via a monthly payment automatically deducted from a credit card. Enterprises have been using hosted applications for learning software for over a decade. Software as a service (SaaS) is one type of computing that is usually in the cloud and delivers a single application through the browser to thousands of customers using a multitenant architecture. The biggest change in attitude towards the cloud has come over time as Chief Technology Officers realize they do not have to maintain software and services within their own buildings and can maintain the same control via renting the software and server capacity. Alternatively, the CTO realizes their kingdoms are at capacity and welcome departments outsourcing to the cloud. Think– enterprise sales tracking installed on each salesperson's desktop containing a copy of ACT or a ten-dollar a month bill for each salesperson's online login to www.salesforce.com.

Therefore, it goes for distance education using the Internet. Advanced Learning Management Systems now also come with services attached. Often administrative support and consulting services are included on the use of the software, allowing the customer to build corporate eLearning viability and online education business offerings.

Learning Services: Delivery of Learning

A learning management system (LMS) delivered via the cloud is generally a web application seamlessly delivered over the Internet, accessible from anywhere in the world. It is hosted on servers at a third parties' data center. The use of the Learning Management System is rented. Advantages for the enterprise are that the software is updated frequently, and does not have to be maintained by the customer. The application is essentially "version-less" in the customer's mind since only one active code release exists. Usually the LMS SaaS provides updates on a quarterly or bi-yearly basis. The using enterprise does not have to purchase hardware or people to operate / set up the servers. Moreover, during peak usage the cloud service increases capacity to service more users. Sophisticated clouds will automatically spawn virtual services to meet increasing demand. Some purists claim that this capacity to automatically spawn virtual services is a key part of being a cloud service. Others use the term more loosely.

[21]Research Scholar, CMJ University, Meghalaya, India, prachimanoj8@gmail.com
[22] Associate Professor, Department of Computer Science, Manav Bharti University, Himachal Pradesh, India, parveenritika@gmail.com

Learning Objects / Modules

The e-Learning pundits are talking about for 2012 is authoring content in the cloud. The cloud provides the capability for collaborative development tools for creating, reviewing, and publishing interactive tutorials, assessments, and learning objects. Typically, the control of authoring e-Learning content rested with individuals working with specific authoring software installed on desktops. Project managers looking for efficiency and repeatability have longed for online systems that allowed for distributed workflow that is scalable. Imagine an online system where subject matter experts can review module pages anytime/anywhere and comment in context and where comments are captured in a database. The pundits are saying that if websites can be built on a "what you see is what you get" model so should it be for eLearning content.

CONCERNS BEFORE DEPLOYMENT

As an m-Learning manager, it is imperative to make sure you are getting the advantages of the cloud when a company uses the buzzword in their product promotion. Perhaps some of the biggest concerns hidden in the haze of the cloud's popularity are its ability to deliver on the promise of redundancy, scalability, and security. Where these virtual servers are housed is a legitimate question to ask. There are no guarantees. Amazon had a 2011 publicized service hiccup due to human error and Sony has been hacked. Known security issues exist with larger brands, whereas many good SaaS providers have impeccable records. So go figure. Here are few topics to ask questions about:

Redundancy and Scalability

The key to redundancy is to design an architecture that does not have any single point of failure. A cloud computing system must make multiple copies of client information and store it on other devices and transfer workloads for easier information retrieval or in case of a break down. Redundancy enables the central server to access backup machines to retrieve data that otherwise would be unreachable. The redundancy associated with clouds is not always a given, but it is easier as a side effect of this structure. Since most of the cloud serving a site should be made out of disposable machines (since it's ideal to shut them down, start them back up, rebuild them, clone them frequently) the loss of one or more virtual machines due to a software or hardware issue is less of a problem than it would be otherwise. This does not apply to all, but to most.

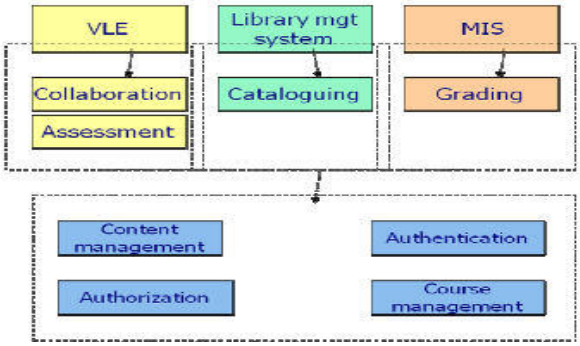
Security

The security of the software running on the cloud is up to whoever is managing it. The security benefits of running in the cloud at a reputable data center must meet certain physical location security measures to accommodate; like HIPPA or other guidelines. Some data centers are certified (SSAE 16) and must submit to security audits. Regularly performing a security analysis is vital to the security of any network. It is the only way to ensure that firewalls and access controls are properly configured and that server updates have been applied. Consider the importance of both physical and electronic security escort-only physical security, alarm system, video surveillance, motion detectors and glass break detectors, and dedicated network security experts. Ask your cloud provider if the data center is certified. All of this can be thrown out of the window when talking about many uses of the cloud. It is a big buzzword/ marketing term and is thrown around a lot. In some cases, it means that someone has placed one or more constantly running virtual machines on a provider, but it is not anything new. It just has a new name. As an eLearning manager looking to deliver, author and store content in the cloud, you should be asking questions about auto redundancy, scalability and security. A reputable company with a handle on these important aspects of cloud computing is much more reassuring than just a brand name.

TRADITIONAL E-LEARNING AND CLOUD BASED M-LEARNING

e-Learning is an Internet-based learning process, using Internet technology to design, implement, select, manage, support and extend learning, which will not replace traditional education methods, but will greatly improve the efficiency of education. As e-Learning has a lot of advantages like flexibility, diversity, measurement, opening and so on, it will become a primary way for learning in the new century as in Figure 1

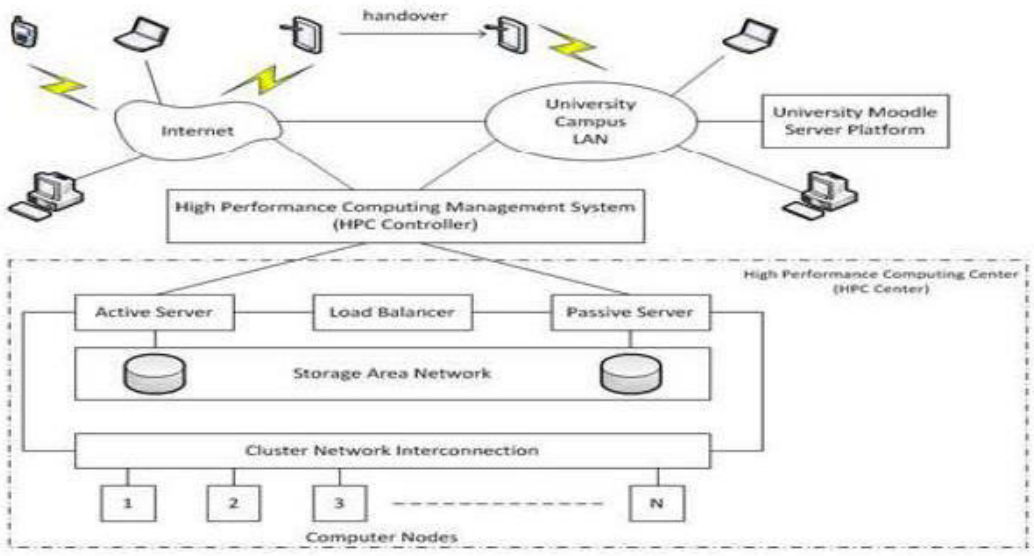
Figure-1: Traditional E-Learning



Sources: Authors Compilation

This paper is going to propose an innovative e-learning ecosystem based on cloud computing and Web 2.0 technologies. The paper analyses the most important cloud-based services provided by public cloud computing environments such as Google App Engine, Amazon Elastic Compute Cloud (EC2) or Windows Azure, and highlights the advantages of deploying E-Learning 2.0 applications for such an infrastructure. The authors also identified the benefits of cloud-based E-Learning 2.0 applications (scalability, feasibility, or availability) and underlined the enhancements regarding the cost and risk management. Our proposed system primarily is composed of different cloud partners, local servers and cloud central system. The architecture is depicted in Figure2.

Figure-2: Cloud Based E-Learning and M-Learning Model

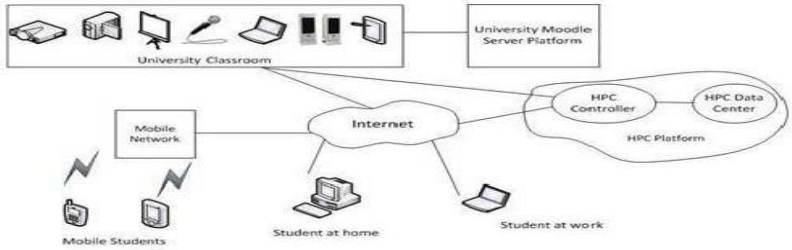


Sources: Authors Compilation

According to our proposed architecture each individual PC act as a cloud partner, which offers the necessary resources to the cloud system from its available resources. However, each of these individual PC is the property of a particular educational institute or University study center whereas all these partners or users owned those tablet PCs like "Aakash" from the budget sanctioned by the government for that particular institute or University. There is a local server associated with individual study center of an institute who monitors everything ranging from PC status to individual requests for that institute. The users associated with a particular local server submit their request to the cloud via the local server. The local server collects the entire request from the clients in its domain within a specific time and forward those request after verification. In addition, some providers have the agreement with the cloud system and offers different services to the user.

The proposed e-learning cloud architecture can be divided into the following layers: Infrastructure layer as a dynamic and scalable physical host pool, software resource layer that offers a unified interface for e-learning developers, resource management layer that achieves loose coupling of software and hardware resources, service layer, containing three levels of services (software as a service, platform as a service and infrastructure as a service), application layer that provides with content production, content delivery, virtual laboratory, collaborative learning, assessment and management features. A Infrastructure layer is composed of information infrastructure and teaching resources. Information infrastructure contains Internet / Intranet, system software, information management system and some common software and hardware; teaching resources is accumulated mainly in traditional teaching model and distributed in different departments and domain. This layer is located in the lowest level of cloud service middleware, the basic computing power like physical memory, CPU, memory is provided by the layer. By virtualization technology, physical server, storage and network form virtualization group for being called by upper software platform. The physical host pool is dynamic and scalable, new physical host can be added in order to enhance physical computing power for cloud middleware services.

Figure-3: Cloud Supported m-Learning Environment



Sources: Authors Compilation

BENEFITS FROM ARCHITECTURE

- Powerful computing and storage capacity: Cloud based E-learning architecture locates the computing and data in a large number of distributed computers, the sea of clouds in the tens of thousands of computers to provide powerful computing power and huge data storage space, puts the "cloud" as a service available to students via the Internet.

- High availability: Through the integration of mass storage and high-performance computing power, this system can provide a higher quality of service. Cloud computing system can automatically detect the node failure and exclude it, do not affect the normal operation of the system.

- High security: In the cloud-computing model, data is storied intensively. Relying on one or more data center, the managers manage the unified data, allocate the resources, balance load, deploy the software, control security, and do the reliable real time monitoring, thus guarantee the users' data security to the greatest possible degree.

- Virtualization: Virtualization is the most important characteristics of this type of architecture. Each application deployment environment and physical platform is not related. It is managed, expensed, migrated, and backup through virtualization platform. It put the underlying hardware, including servers, storage and networking equipment, comprehensive virtualization, in order to build a resources pool of shared, distributed on-demand.

The major advantage of the proposal is that it aims at providing easy access to costly software running on high performance processors to rural students at institutions, which lack considerable facilities. Considerable investment would be required to implement this architecture, but the benefits would easily justify the cost.

CONCLUSIONS

The e-Learning model cannot completely replace teachers; it is only an updating for technology, concepts and tools, giving new content, concepts and methods for education, so the roles of teachers cannot be replaced. The teachers will still play leading roles and participate in developing and making use of e-learning cloud. The blended learning strategy should improve the educational act. Moreover, the interactive content and virtual collaboration guarantee a high retention factor. On the other hand, E-learning cloud is a migration of cloud computing technology in the field of e-learning, which is a future e-learning infrastructure, including all the necessary hardware and software computing resources engaging in e-learning. After these computing resources are virtualized, they can be afforded in the form of services for educational institutions, students and businesses to rent computing resources.

Present economic situation will force different educational institutions and organizations to consider adopting a cloud solution. Universities have begun to adhere to this initiative and there are proofs that indicate significant decreasing of expenses due to the implementation of cloud solutions. The aim of our work was to identify an architecture, which will be using Cloud Computing within school level or higher education. Mainly, we have considered the benefits of cloud architecture. Future research will include a study regarding the attitude and strategy for migration to the proposed architecture based on clouds.

REFERENCES

1. Armbrust, M., Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., …Zaharia, M. (2009, February 10). ***Above the Clouds: A Berkeley View of Cloud computing*** (Technical Report No. UCB/EECS-2009-28). USA: University of California at Berkley.

2. Vaquero, L., Merino, L. & Caceres, J. (2009). A break in the clouds: towards a cloud definition. ***SIGCOMM Comp. Communications Review,*** 39, 50-55.

3. Youseff, L., Butrico, M., & Silva, Da D. (2008). Toward a unified ontology of cloud computing. ***Grid Computing Environments Workshop (GCE '08),*** pp. 1-10.

4. Mell, P., & Grance, T. (2009). ***The NIST definition of cloud computing***. National Institute of Standards and Technology.

5. Shufen, Zhang, Shuai, Zhang, Xuebin, Chen, & Shangzhuo, Wu. (2010). Analysis and Research of Cloud Computing System. In ***Instance 2010 Second International Conference on Future Networks,*** (pp. 88-92).

6. Xu, Lei, Xin, Zhe, Ma, Shaowu, & Tang, Xiongyan. (2009). Cloud computing and services platform construction of telecom operator. Broadband Network & Multimedia Technology, 2009. IC-BNMT '09. ***2nd IEEE International Conference on Digital Object Identifier*** (pp. 864-867).

7. Thomas, Dietinger. ***Aspects of e-Learning Environments*** (Doctoral Dissertation). Graz University of Technology. Retrieved on January 12, 2008 from http://www.iicm.tugraz.ac.at/thesis/tdieting_diss.doc

8. ***Ministry of Education Government of People's Republic Bangladesh.*** Retrieved from http://www.moedu.gov.bd/.

9. Wei, Y., & Rong, Y. (2004, November). ***Research of an E-learning System Model Based on Agent.*** Computer Engineering and Applications, 156-158.

10. Gladun, A., Rogushina, J., Garcı′a-Sanchez, F., Martı′nez-Be′jar, R., & Toma′s Ferna′ndez-Breis, J. (2009). An application of intelligent techniques and semantic web technologies in e-learning environments. ***Expert Systems with Applications,*** 36, 922-1931.

11. Li, Y., Yang, S., Jiang, J., & Shi, M. (2006). Build grid-enabled large-scale collaboration environment in e-learning grid. ***Expert Systems with Applications,*** 31, 742-754.

12. Chengyun, Z, (2010, May). ***Cloud Security: The security risks of cloud computing, models and strategies.*** Programmer, 71-73.

13. Hayes, B. (2008). Cloud computing. ***Comm. Acm,*** 51(7), 9-11.

14. Tuncay, E. (2010). ***Effective use of Cloud computing in educational institutions.*** Procedia Social Behavioral Sciences, 938-942.

15. Buyya, R., Yeo, S. C., & Venugopal, S. (2009). Market oriented Cloud computing: Vision, hype, and reality of delivering IT services as computing utilities. ***10th Ieee Int. Conf. High Performance Comput. Comm***. (pp. 5–13).

16. Lijun, M., Chan, K. W., & Tse, H. T. (2008). A tale of Clouds: Paradigm comparisons and some thoughts on research issues. ***Ieee Asia-pasific Services Comput. Conf., Apscca08*** (pp. 464-469).

17. Praveena, K. & Betsy, T. (2009). Application of Cloud Computing in Academia. ***Iup J. Syst. Management,*** 7(3), 50-54.

18. Delic, A. K., & Riley, A. J. (2009). Enterprise Knowledge Clouds. In ***Next Generation Km Syst. Int. Conf. Inform., Process, Knowledge Management*** (pp. 49-53), Cancun, Mexico.

*****

PEZZOTTAITE JOURNALS MESSAGE TO AUTHORS

We require that, prior to publication; authors make warranties to these effects when signing their Agreements.

An author must not submit a manuscript to more than one journal simultaneously, nor should an author submit previously published work, nor work which is based in substance on previously published work.

An author should present an accurate account of research performed and an objective discussion of its significance, and present sufficient detail and reference to public sources of information so to permit the author's peers to repeat the work.

An author must cite all relevant publications. Information obtained privately, as in conversation, correspondence, or discussion with third parties, should not be used or reported in the author's work unless fully cited, and with the permission of that third party.

An author must make available all requisite formal and documented ethical approval from an appropriate research ethics committee using humans or human tissue, including evidence of anonymisation and informed consent from the client (s) or patient (s) studied.

An author must follow national and international procedures that govern the ethics of work done on animals.

An author must avoid making defamatory statements in submitted articles which could be construed as impugning any person's reputation, for example, making allegations of dishonesty or sharp practice, plagiarism, or misrepresentation; or in any way attacking a person's integrity or competence.

An author must ensure all named co-authors consent to publication and being named as a co-author, and, equally, that all those persons who have made significant scientific or literary contributions to the work reported are named as co-authors.

Additionally, the author understands that co-authors are bound by these same principles.

(sd/-)
(Editor-In-Chief)

CUSTOMER AWARENESS OF ONLINE BANKING: A THEORITICAL PERSPECTIVE

S. Sankar[23] K. Maran[24]

ABSTRACT

The purpose of the paper to measure the awareness of online banking and effectiveness of the banking system that enables bank customers to access accounts and general information on bank products and services through a personal computer or other intelligent device; as the banking products and services can include wholesale products for corporate customers as well as retail and fiduciary products for consumers. Ultimately, the products and services obtained through Online banking may mirror products and services offered through other bank delivery channels.

KEYWORDS

Online Banking, Computer, Awareness, Services, Products, Security etc.

INTRODUCTION

In the present era of globalization and privatization, the only word presumed to be constant is the word 'change'. Globalisation has reached all the nooks and cranny of the world, showing positive effects and remarkable developments and innovation especially in the communication and technology field, it can be seen that the tricks of the trade are changing very fast. The development of communication and technology has definitely made the world much smaller and brought it within the boundary of our living room. However, the most beneficial has been the evolution of internet. Internet and its allied services have really done wonders in almost every aspect of life. The commercial world has benefited the most from the rise of the internet. it can be said that the internet covers almost all the aspects of life. The online aspects have really made the financial world much better place to be it. Among the various financial sub-sectors, the commercial banks are the most important. And the introduction of internet in the daily operation of such financial banks have made the transaction all the more beneficial for all the concerned parties be it the valued customers, the commercial bank, government or the monetary authority be it the media.

The online/Internet Banking Internet banking is a time saver, instead of going into the bank for transactions, you can easily do it on the internet. The internet banking is the call of the hour. Because of the ever-increasing competition, it is always accepted by all the business leaders that it is difficult to get a new customer than to retain an existing one. Therefore, duty lies on the service provider that has the best of the services provided to the existing customers so that they do not shift their business to other new entrant or an existing player. It is all the same for the banking industry. Even if the leading banks of the India Nationalized and Private sectors banks the latest technologies and upgrade them, it is almost certain that they would lose their customer base. The leading four banks of the United Kingdom have the following facilities within the range of the online banking i.e. the internet banking and the mobile banking.

CURRENT ONLINE/INTERNET BANKING OFFERINGS IN INDIAN BANKS

The leading bank tries to put emphasis on the point that the customer should be able to keep track of their money and should have the opportunity to withdraw when ever and where ever he feels as far as it is his money. The bank rightly says that the online banking also enables the customer to check statements, transfer funds and pay the bills. Indian banks depend upon three important pillars to promote online banking among its customers namely convenience, easy operations and secured transactions.

In its website, the bank specifies in clear terms the steps through which the valued customer can register for the online banking services along with generation of five (5) digits pass code or PIN sentry, which would ultimately enable the customer to get the full access of the online banking. Pertaining with the fact that a major part of the customers might not be yet conversant with the online banking system, Indian banks have a special section in its website containing the detailed demonstration, online banking system for business needs and most importantly the online banking guarantee.

ONLINE BANKING

The banks have brought in whole lot of services under the department of internet banking. Some of the basic facilities are listed below:

- Search Statements - This facility enables the customer to search and verify statements by payment type, date or amount since July of 2002. It also facilitates him to download the statements into spreadsheets or any of the financial software packages.
- Stop paper statements - This feature serves as the corporate social responsibility as the usage of paper reduces as statements can be viewed electronically. In addition, the features allow the customer to get the duplicate copy whenever required.
- Real time balances - The balance can be checked on the transaction basis and a track can always be kept on the account.

[23]Research Scholar, Sri Sairam Institute of Management, Tamil Nadu, India, aathiya2002@yahoo.co.in
[24]Director, Management Studies, Sri Sairam Institute of Management, Tamil Nadu, India, maran.mba65@gmail.com

- Moving money and paying bills - The money can be transferred within the bank accounts instantly and also bills can be paid apart from paying to families and friends through faster services.
- Manage bills online - it includes receiving bills online, getting email alerts and paying as and when the customer wishes.
- *All the Bank accounts can be checked and the credit limit of the credit can be fixed.
- Getting Free Tax alerts - The customer has the option to register free tax alerts; he can also order cheque and paying books through the online mode.
- Managing direct debits and standing orders - The bank also allows setting up, changing and cancelling standing orders and managing direct debits.
- Hassle free application for products and services - The customer can apply new or increased overdraft as well as loan, opening online savings account and can upgrade current account.
- Ordering travelers' cheque as well as foreign currency - This feature is very helpful to the international customers as it initiates commission free online travel money.

In case of mobile banking tax alerts are made available along with the facilities like checking of balance and mini-statements and transferring of funds. Also, the bank allows ordering debit card, credit card, cheque book and paying-in book through mobile banking.

CUSTOMER ACCEPTANCE OF ONLINE BANKING

Researcher on customer behaviour and acceptance of online banking revealed that there are several factors that affects a customer behaviour towards online banking, such as the demographic factor, likewise prior experience to new technology and computers also influence the behaviour towards online banking. A typical user of online banking in a Finnish market, highly educated, relatively young and wealthy with a good knowledge of computers especially the internet. The outcome of the research suggested that demographic factor have an effect on his online banking attitude. The younger people/customers value the convenience or the time saving potential of online banking more than the older generation.

However, the avoidance of meeting a bank official is taken as important by the younger generation unlike the older generation. The less privileged people of the society will prefer the Face-to-Face banking, unlike the rich, which will prefer to use online banking. Nevertheless, the use of online banking has been argued that demographic, geographic, social- economic and psychographic variable have not been generally accepted as a viable predictors in determining the purchasing attitude/behaviour in the financial services by past and recent studies. Instead segmenting the customer in bank marketing by expected benefits and attitude. These authors used the cluster analysis to separate the customers into four groups, that is, "the transaction oriented group, who have a relatively strong technology but their attitude to information is weak, "the generally interested": they have a positive technology with a strong attitude towards information, "the service oriented group: these set of people have a weak attitude towards technology and information. Lastly, the "technology opposed" group: these set of people have a very strong attitude towards information but weak attitude to technology. But the above study was recently contradicted by a study done by Sarel and Marmorstein (2003) which revealed that the income of a household and the level of education significantly affects the acceptance of online banking.

RESEARCH OBJECTIVES

The acceptance of online banking in the India has had a tremendous change in the customers and the banking sector, hence it is essential to investigate the market and deduce the true customers and the demographics. Therefore, the objectives of this research were:

- To understand the perception of customer to online banking.
- To truly identify the target customer of online banking.
- To equate attitudes of users of online banking.

GENERAL FINDINGS

Some previously researched reports were able to deduce that the internet is about the biggest thing to hit the world, and are still making an impact, by all indications, it will continue to do in the next couple of years. To align the life with this biggest and the most amazing as well as technological upsurge, almost all services and products all across the life has been using internet in its favour. The increasing number of internet user's proves the fact that the internet just cannot be ignored. There has been steady rise of the electronic customers for the banning majors of India and as more and more of the banks get into it along with large number of customers, they can reduce cost in terms of work force employed and increase profitability thus creating value for the shareholders. In addition, the fact is that still large number of banking persons solely depend on branch banking and the banking companies should try to tap them into online banking by making them understand the various benefits of the system. Nevertheless, proving all the misconceptions wrong and to reap the benefits of the online banking industry, the numbers of persons using online banking for various reasons are also on the rise. All the above suggest that there is a huge boom in the online banking system and the customers are increasingly being hooked to it because of its superb features.

EVALUATION OF BANKS WEBSITES

Almost all of the major banking groups all across the globe has divulged in to online banking understanding the call of the hour The prime reasons of such enormous popularity of the online banking system has been the vital factors of cost savings (by

reducing the man power), increased customer base (more and more people are getting themselves under the banking system and it is almost impossible to serve all of them manually at the branches), mass - customization (the individual can be targeted personally and his own choices can be addressed in place of standardization), increased innovation and development of non core business and less expensive marketing campaigns.

However, banks websites were critically examined in terms of Security, Speed, Content and Design and Navigation:

- **Speed**
  Speed is very important in the cyber world. One of the prime reasons of the success of internet is its fastness and now if the web service initiated by the bank is very slow, the internet community user would be utterly disappointed and it might even so happen that they turn away from the service or even changes their banking partners.

- **Content and Design**
  The content and design influences user satisfaction. In their study of web user satisfaction, Product information content, Amount of product information, Product information format, Language(s) and Layout features were characteristics found to have great influence on users. Some bank website were found to have plain design and plain graphics, while come were with attractive design and graphics as they were used for marketing and advertisements for the banks.

- **Navigation**
  Navigation depends upon lot many factors. It comprises of the hyperlinks, web pages, homepages, and amount of hyperlinks, hyperlink encoding and personal contact possibilities among others, which would influence the satisfaction of the use. According to one of the respondent on navigation of the banks websites, she revealed that: "I find it easy to use as in, everything is in bold writing, everything in different category and to get to the internet banking, you can just type in from Google and it would come up".

- **Security**
  This could be seen as the important issue affecting the existing user and intending users because there have been many experiences where the password of customers were been hacked and money taken out of their account(s). However, with the introduction of advanced technology, sophisticated software's has been introduced into various banks website. This software's are to protect user's passwords and users Ids from hackers. Encryption technology is also one of the security features in all banks websites. In addition, most bank websites request for some digits/details (usually sent directly to the customer by the bank) from the user before giving them the access to the page where their password is required. Although, all facilities are still in their early stage which means there is stillroom for more development in the security features in banks websites.

INTERNET BANKING: RELEVANCE IN A CHANGING WORLD

Surprising, but true - Internet-based activity is not the preserve of the young "digital native" generation alone. A 2008 survey says that Generation X (those born between 1965 and 1976) uses Internet banking significantly more than any other demographic segment, with two thirds of Internet users in this age group banking online. Gen X users have also professed their preference for applications such as Facebook, to share, connect and be part of a larger community. This is some irony in this, since online banking, as we know it today, offers minimal interactivity. Unlike in a branch, where the comfort of two way interaction facilitates the consummation of a variety of transactions, the one way street of e-banking has only managed to enable the more routine tasks, such as balance enquiry or funds transfer. It's not hard to put two and two together. A clear opportunity exists for banks that can transform today's passive Internet banking offering into one that provides a more widespread and interactive customer experience. It is therefore imperative that banks transform their online offering, such that it matches the new expectations of customers. Moreover, Internet banking must journey to popular online customer hangouts, rather than wait for customers to come to it. There are clear indications that the shift towards a "next generation" online banking environment has already been set in motion. It is only a matter of time before these trends become the norm.

LEVERAGING OF SOCIAL NETWORKS

Forward thinking banks are leveraging existing social networks on external sites to increase their visibility among interested groups. They are also deploying social software technology on their own sites to engage the same communities in two way discussions. Thus, their Internet banking has assumed a more pervasive persona - customers are engaging with the bank, along with its products and services even when they're not actually transacting online. Heightened visibility apart, banks can gain tremendous customer insight from such unstructured, informal interactions. For example, a discussion on the uncertain financial future among a group of 18 to 25 year olds could be a signal to banks to offer long term investment products to a segment that was previously not considered a target. Going one step further, a positive buzz around a newly launched service can create valuable word-of-mouth advertising for the business.

COLLABORATING THROUGH WEB 2.0

The collaborative aspect of Web 2.0 applications has enabled banks to draw customers inside their fold more than ever before. Traditional methods such as focus group discussions or market research suffer from the disadvantages of high cost, limited scope and potential to introduce bias. Feedback forms merely serve as a post-mortem. In contrast, Web 2.0 has the ability to carry a vast

audience along right from the start, and continue to do so perpetually. Thus, an interested community of prospects and customers participate in co-creating products and services, which can fulfill their expectations. The pervasiveness of Web 2.0 enables delivery of e-banking across multiple online locations and web-based gadgets such as Yahoo! Widgets, Windows Live or the iPhone. This means next generation online banking customers will enjoy heightened access and convenience A New York based firm of analysts found that 15% of the 70 banks tracked by them had adopted Web 2.0, a number of them having done so within the last 12 months.

PERSONALISATION OF ONLINE BANKING

e-banking divides customers into very large, heterogeneous groups - typically, corporate, retail or SME, with one type of Internet banking page for each. That is in sharp contradiction to how banking organizations would like to view their clientele. Banks are moving towards customer-specificity, almost viewing each client as a "segment of one", across other channels, and online banking is set to follow suit. For instance, a specific home page for home loan customers and another for private banking clients could well be a possibility in future. Money Monitor from Yes Bank allows customers to choose their landing page - for example, they can set "all transactions", "net worth" or "portfolio" as their default view. Other features include the ability to categories transactions as per customers' convenience and the printing of custom reports.

EMPOWERMENT ONLINE

Beyond doubt, Internet banking has created a more informed, empowered class of customers. This is set to climb to the next level once customers are allowed to proactively participate in many more transaction-related processes. The Internet has already made it possible for customers to compare product loan offerings, simulate financial scenarios and design custom retirement portfolios. Going forward, they would be able to consummate related transactions - which means, after comparing interest rates, they could originate a loan online, and once secured, they can begin to repay it online as well.

PORTALISATION

The emergence of Web 2.0 technology coupled with banks' desire to personalize their e-banking to the highest degree is likely to result in "portalisation" of Internet banking. The idea of banking customers being able to create their own spaces online, filled with all that is relevant to them, is not that far-fetched. Customers can personalize their Internet banking page to reflect the positions of multiple accounts across different banks; they could include their credit card information, subscribe to their favourite financial news, consolidate their physical assets position, share their experiences with a group and do more - all from one "place".

Money Monitor enables customers to add multiple "accounts" (from a choice of 9,000) to their page. Accounts could be savings or loan accounts with major Indian banks, or those with utilities providers, credit card companies, brokerage firms and even frequent flyer programs. Users can customize their pages as described earlier. As banks seek to develop their Internet banking vision for the future, in parallel, they will also need to address the key issues of security and "due defense". While it is every marketer's dream to have customers work as ambassadors, adequate precaution must be taken to prevent the proliferation of malicious or spurious publicity. Therefore, before an individual is allowed to participate in a networking forum, he or she must have built up a favorable track record with the bank. The individual must be a recognized customer of the bank, having used a minimum number of products over a reasonable length of time. Qualitative information about the person's interaction with the bank's support staff (for example frequency and type of calls made to their call centre, outcome of such interaction and so on) may be invaluable in profiling the "right" type of customer who can be recruited as a possible advocate.

Collaborative Web 2.0 applications may necessitate opening up banks' websites to outside technology and information exchange with third party sites, raising the spectra of data and infrastructure security. A robust mechanism of checks and balances must be built to ensure that the third party sites are secure, appropriately certified and pose no threat to the home banks' sites. Likewise, before a third party widget is allowed to be brought on to a site, it must have passed through stringent security control. Due diligence must be exercised before permitting users to place a link to another site to guard against the possibility of inadvertent download of malicious software, which could, in the worst case, even result in phishing originating from the banks' sites. It is equally important for a bank to guard its customers against invasion of privacy, data theft or misuse. The concept of portalisation envisages deploying technology to bring information from other banks' or financial service providers' websites into the home bank's site. The home bank must ensure that its customers' personal or transaction related information, which may be shared with the other providers, is not susceptible to leakage or outright misuse.

Banks will do well to partner with an Internet banking solution provider which has not only the expertise to translate their vision into a cutting edge e-banking experience for the user, but also the foresight to define boundaries for safety. With security concerns adequately addressed, next generation Internet banking is full of exciting possibilities. Banks that seize the opportunity may find that Internet banking can become a means of differentiating themselves from competitors, rather than a mere cost cutting tool. Clearly, providing a more powerful and interactive e-banking experience, is the way forward. Finacle consumer e-banking solutions are proven Internet banking and mobile solution for retail banking customers.

CONCLUSIONS

There is no doubt that the online banking system is one of the most happening and beneficial services that has really changed the tricks for one of the most traditional industry i.e. banking. The trends have always favoured the fact that the number of online banking users are on the rise. The leading Indian commercial banks have played pivotal role in the upsurge of the system in the

India. However, the bottom line fact remains that he online banking system can easily accommodate more of users and serve them with hassle free transactions with utmost speed and safety. The banks should encourage such practices and initiate measures, which are beneficial to the customers as well as the shareholders.

REFERENCES

1. Rolf, Oppliger. (1996). ***Authentication systems for Secure networks.*** Retrieved from www.artech-house.com, rolf.oppliger@esecurity.ch, Artech House.
2. ***Packet Magazine-*** Third Quarter 2000 Issue – CISCO systems (packet@cisco.com).
3. ***A step-by-step guide for secure online commerce,*** Verisign. Retrieved from www.versign.com
4. A hacker's Guide to protecting Your Internet Site and Network –Anonymous. ***Maximum security.*** Retrieved from www.ods.com.ua/win/eng/security/Max_Security
5. CERT® Security Improvement Modules. ***CERT.*** Retrieved from www.cert.org/security_improvement
6. ***Electronic Banking Safety and Soundness Examination Procedures***. Federal deposit insurance corporation division of supervision. Retrieved from www.fdic.gov/regulations
7. (1999). ***The Information Technology Bill***. Retrieved from www.mit.gov.in/bill99.htm
8. ***E Banking: Risks And Responses Carol Sergeant Director.*** Banks & Buildings Societies Financial Services Authority. Retrieved from www.fsa.gov.uk
9. (1996, October). ***Bank for International Settlements- Implications for Central Banks of the Development of Electronic Money.*** Retrieved from www.bis.org/publ/Disp01.pdf
10. ***Enabling E-Commerce in India.*** Retrieved from www.giic.org
11. ***Business models for electronic markets***. CommerceNet. Retrieved from www.commerce.net
12. (1998). ***E-Commerce: A white Paper***. Keith Hazelton University of Wisconsin- Madison.
13. (1999, October). ***Comptroller of the Currency*** (Internet Banking Comptroller's Handbook).

*****

THESIS ABSTRACT

ON

'THE DATA WAREHOUSE DESIGN PROBLEM THROUGH A SCHEMA TRANSFORMATION APPROACH'

BY

V. Mani Sarma[25] P. Premchand[26]

ABSTRACT

An operational database is a shared repository of data. The OLTP's are used for day-to-day transaction processing and do not fulfill the out of the ordering tasks of information processing and data analysis. A new design methodology is required to map the OLTP DB's data into MD data model. The MD data model is a logical data model comes from the enterprises, they have amount of data, but they cannot access it. They use the same search patterns day to day and data usually is not prepared for an easy queries. The MD model represents a measure that depends on a set of dimensions, which provides the context for the measure it can be represented by hyper cube, that are the graphic way and conceptual model or by the schemas in a logical model. Previously DWH design and development applied ad-hoc methodologies. An increasing number of organizations are implementing data warehouses to strengthen their decision support systems. This comes with the challenges of the population and the periodic update of data warehouses. In this thesis, we present a tool that provides users with features to create a warehouse database and transform structures of the source database into structures for the warehouse database. It is highly interactive, easy to use, and hides the underlying complexity of manual SQL code generation from its users. Attributes from source tables can be mapped into new attributes in the warehouse database tables using aggregate functions. Then, relevant data is automatically transported from the source database to the newly created warehouse. The tool thus integrates warehouse creation, schema mapping and data population into a single general-purpose tool.

This tool has been designed as a component of the framework for an automated data warehouse. Users of this framework are the database administrators, who will also be able to synchronize updates of multiple copies of the data warehouse. Warehouse images that need to be updated are taken offline and applications that need to access the data warehouse can now access any of the other image warehouses. The Switching Application built into this framework switches between databases in a way that is very transparent to applications so that they do not realize existence of multiple copies of the data warehouse.

KEYWORDS

Data, Warehouse, Design, Schema Evolution, Schema Transformation, Relational DW, DW Designs Trace, MD Model, OLAP, MDBE, SAMDSG etc.

INTRODUCTION

Database creation is a complex task, involves tuning many parameters, and can be done using any database creation. If many databases need to be created, the database creation wizard will need to be used repeatedly for each new database. This can be cumbersome especially when only a minimum number of essential parameters differ for each database.

Data warehouses are databases that are loaded with subsets of relevant data from a source database. These warehouses may contain informational data extracted from operational data in the source database. The tables in warehouse databases are based on the tables from the source database. Hence, it is essential to transform structures of the source database into structures for the warehouse. Nowadays, this is done by manually exploring and creating such a mapping. This process is both tedious and time-consuming. In addition, users need to be technically trained to perform this task. There are a few other shortcomings in the present system. In the warehouse schema, users may add new attributes to tables; these new attributes are the aggregates of the attributes of the master database. As a result, when data is copied from the master database to the warehouse database, data for these aggregate functions need to be computed at run-time during update, causing more delay. When this update is in progress, applications accessing the warehouse will not get access to accurate data, leading to lack of synchronization. MDBE follows a classical approach, in which the end-user requirements are well known beforehand. This approach benefits from the knowledge captured in the data sources, but guides the design task according to requirements and consequently, it is able to work and handle semantically poorer data sources. In other words, providing high-quality end-user requirements, we can guide the process from the knowledge they contain, and overcome the fact of disposing of bad quality (from a semantically point of view) data sources.

[25]Associate Professor, Nalla Malla Reddy Engineering College, Andhra Pradesh, India, India, manisharma.vittapu@gmail.com
[26] Dean, Department of Computer Science and Engineering, Osmania University, Andhra Pradesh, India, drpreamchand_p@yahoo.com

OBJECTIVES OF STUDY

These problems form the basis and the motivation for this thesis. The Semi-Automatic Multidimensional Schema Generation (SAMDSG) tool is works towards providing an interface to accept required information from users to generate a new multidimensional database and creates an empty data warehouse. For a given source database, the SAMDSG tool aims at arriving at an appropriate mapping to create a warehouse structure. After a mapping has been formalized, tables for the new warehouse are created. Then, relevant data is automatically transformed from the source database to the newly created warehouse.

A framework has been built to facilitate automatic updates of data warehouses. It has been designed in a way that the there can be multiple copies of the warehouse database, where each copy is an image of the warehouse database. Copies that need to be updated are taken offline and applications that need to access the warehouse database can now access any of the other image warehouses. The SAMDSG tool also helps to switch application – Image Switcher, switches between databases in a way that is very transparent to applications so that they do not realize existence of multiple warehouse databases.

As a result, using the SAMDSG tool by end user can directly create the desired warehouse schema. A major advantage in using this SAMDSG tool is to automate the SQL script generation for schema creation and data management. The use of such a tool gives the users less time to design DW schema more accurately and efficiently rather than developing the code itself.

CONCEPTS OF AN AUTOMATED DATA WAREHOUSE

A Semi-Automatic Schema Generation Mapping (ASGM) is a tool that works towards providing an interface to accept from users the required information to generate a new database and creates an empty data warehouse. For a given source database, the tool aims at arriving at an appropriate mapping to create a structurally related warehouse. After a mapping has been formalized, tables for the new warehouse are created. Then, relevant data is automatically transported from the source database to the newly created warehouse.

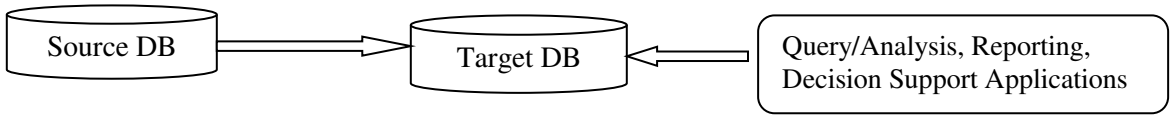
Data Warehouses

Data warehousing is a collection of technologies that support management decision-making. They can be broadly classified as decision support systems. Decision support is a methodology designed to extract information from data and to use this information as a basis for decision-making.

For example, is a component of the Web-based Benchmark data engine? It is a decision-supporting application that extracts data from a warehouse and presents it in visual form. An advantage of manufacturing data in the form of a visual may lead to detection of trends, relationships, exceptions and patterns in the data, if any. Thus, these observations may be a key in the future decision making processes.

Data warehouses contain a wide variety of data that present a coherent picture of business conditions at a single point in time. Informational data is extracted from operational data in the source database and is transformed for end- user decision-making. The development of a data warehouse includes development of systems to extract data from operational systems and the installation of a warehouse database system that provides managers with flexible access to the data. It is aimed at reducing complexity and improving efficiency of data querying. Data warehouses are based on open systems and relational databases. Data warehouses offer organizations the ability to gather and store enterprise information in a single conceptual enterprise repository. Basic data modeling techniques are applied to create relationship associations between individual data elements or data element groups. These associations, or models, often take the form of entity relationship mapping.

Figure-1: Data Warehouse Architecture



Sources: Authors Compilation

The major components are as follows:

Source Databases: In this thesis, data from a source database is migrated to a newly created data warehouse. The master database acts as the source database.

Data Warehouse (Target Database): The target database is the data warehouse that we aim to create and populate. Subset of relevant data and summary data from the source database exists within the data warehouse architecture. It is accessed through queries by desktop applications such as query and analysis, decision support applications and data mining tools.

Decision Support Applications: Decision support applications will use data from the data warehouse. In this thesis, an application such as Chart Visio may act as a Decision Support Application.

Schema Mapping

Data warehousing involves mapping subsets of relevant data from the source database to the target database. The target database schema is designed based on the data that is being transported from the source database. Hence, there is a mapping between the structure of the source database and that of the target database. This mapping is termed as Schema Mapping.

A data warehouse is created autonomously, based on the schema of source database. Schema Mapping is an essential means to transform structures of the source database into structures for the warehouse. This can be done manually, where users can manually explore on creating such a mapping, which can be both tedious and time-consuming. This also assumes that users are technically trained to perform this task.

Automated Data Update in Data Warehouse

Data warehouses contain a wide variety of data that present a coherent picture of business conditions at a single point in time. Informational data is extracted from operational data in the source database, transformed for end-user decision-making and stored in the data warehouse. Since the data in operational databases is continuously changing, it is almost impossible to have real-time replication while maintaining a data warehouse.

Transforming data from the source database to the data warehouse is a slow process. The tables in the warehouse schema may have new attributes, which may be aggregates of the attributes of the master database. During data transfer from the master database to the warehouse database, data for these aggregate functions need to be computed at runtime, causing more delay. At this time, the data warehouse will be unavailable to applications accessing it and these results in lack of synchronization.

Problem Definition

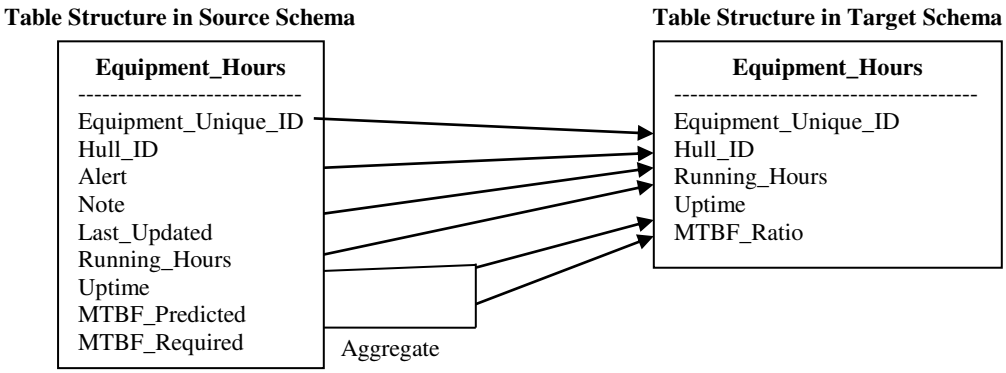
Let us consider the following scenario: Users have a large database and need to store a sub-set of data in a warehouse. The process involved in doing that is:

- Explore the source database and decide what data needs to be represented in the warehouse.
- Create data warehouse by tuning parameters using database creation wizard.
- Form SQL queries to create schema for the newly created data warehouse.
- Form SQL queries to transfer appropriate data from the source database to the data warehouse.
- Periodically manage the update of the data warehouse so that changes in the source database are reflected in the data warehouse.
- Manage multiple images of the data warehouse in order to ensure availability of data warehouse at all times.
- Provide applications with transparent access to multiple images of data warehouse.

This procedure assumes that end-users are familiar with SQL and mandates them to employ other available software to create a data warehouse. Automatic update of the data warehouse needs to be implemented using advanced database concepts. This is time consuming and requires extensive technical support for non-technical users.

Let us consider an example from a sample manufacturing industry database to explain the problem better. (Note: The manufacturing industry database is referred to as the 'source database' in this discussion.) In this scenario, users need to store performance related information of equipments in a data warehouse. Several tasks need to be performed to successfully create such a data warehouse. In the source database, users need to select performance-related data stored in the equipment hour's table, in attributes - Running Hours, Uptime, MTBF Predicted and MTBF Required. In addition, it is essential that the primary key of the Equipment Hours table is a part of the data warehouse. This is presented in figure2.

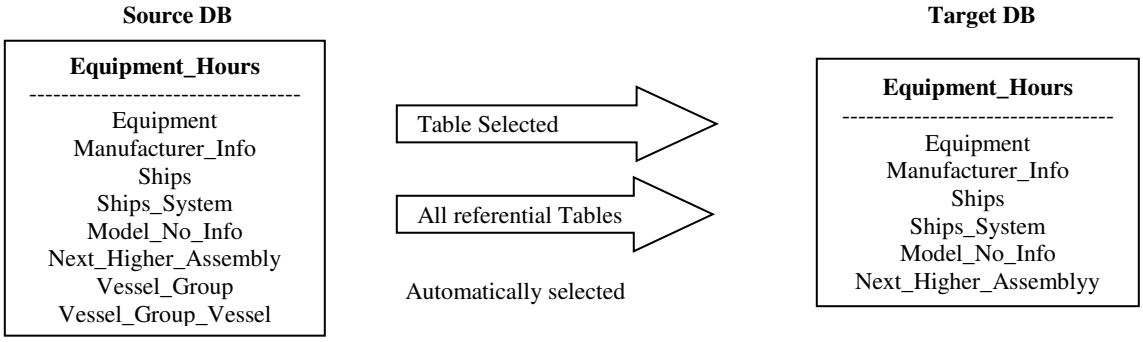
Figure-2: Attribute Selection from Source Table



Sources: Authors Compilation

Users need to create a data warehouse by tuning parameters, using Oracle's database creation wizard. Users need to create the Equipment Hours table in the data warehouse, with only the required attributes and transfer corresponding data. It has to be noted that the Equipment Hours table has several referential integrity constraints, due to which all the parent tables need to be a part of the data warehouse schema. Parent tables of Equipment Hours table are presented in figure3.

Figure-3: Tables Selection from Source Database



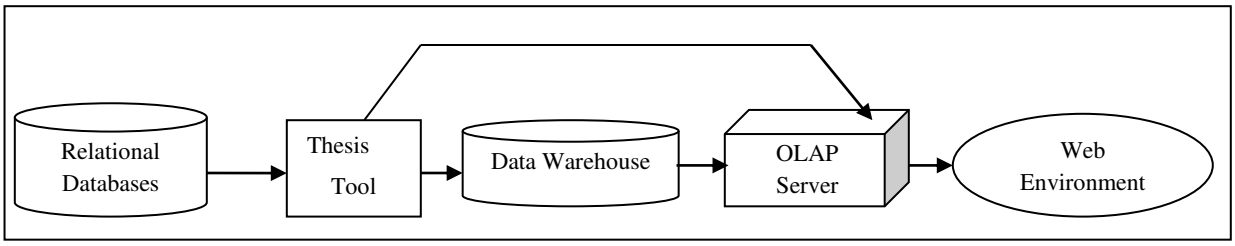
Sources: Authors Compilation

Users may also want to include a new attribute 'MTBF Ratio' that calculates the ratio as an aggregate of MTBF Predicted and MTBF Required. In this example, we have considered mapping only one table from the source database to the data warehouse. The problem grows as the number of tables and attributes to be mapped increases.

THE SOLUTION TO PROBLEM

As a solution to the problems mentioned above, the proposed tool that generates a new data warehouse, performs schema mapping and builds a framework for automatic update of the data warehouse. The proposed tool is to allow users to select, extract, clean, and convert data from source system structures into consistent target warehouse data structures. In addition, the data from the source database is populated into the target database. The data warehouse can be populated on a frequency that meets the organization's needs. The tool navigates users in a sequence of interactive steps and accepts the parameters to create a new data warehouse. For a given source database, the tool helps users in arriving at an appropriate mapping to create a structurally related warehouse. After a mapping has been formalized, tables for the new warehouse are created. Then, relevant data is automatically transformed from the source database to the newly created warehouse. To enable automatic update of the warehouse database, a setup has been built that manages the periodic update of the warehouse. Applications access the data warehouse through an interface that provides a simple-to-use API. Users may create multiple images of the data warehouse using the tool. The support to update all the images is provided in the framework.

Figure-4: System Architecture



Sources: Authors Compilation

ARCHITECTURE OF SYSTEM

System architecture is a vital component of an application design. Architecture translates the logical design of the application to a structure that defines the interaction of entities in the system. The proposed system design in this thesis in order to resolve the problem of the MD database is the one shown in the figure5. This system is composed by different components: a relational database, the thesis tool, a data warehouse, an OLAP server and a web environment. The relational database is a database stored in a database manager. This database manager has to be a PostgreSQL manager. The thesis tool is a program built in JAVA. This tool has to connect to the relational database manager in order to read a database. It has to connect to the data warehouse manager in order to create a new database and introduce data. The thesis tool has to access to the OLAP server directory, in order to store the OLAP schema and the MDX query. The Data warehouse is the database that the thesis tool has created. It is store in a relational database manager. This database manager has to be a PostgreSQL manager. The OLAP server is an open source server belongs to the platform business intelligence, which provides a complete suite to analyses business data. This OLAP server is a module that runs into a Tomcat web server. The Web environment is an access to the data through http channel.

DATABASE CREATION

Database creation is a complex task and involves tuning many parameters. This chapter describes how SagaMap provides a graphical interface to accept the essential parameters to generate script files, which can be executed from the command prompt to create a new database. One of the overall goals of this projects it to be able to create and populate databases for data warehouses. This involves creating a blank database, in which data may be filled in. In this chapter, we discuss how our tool helps towards reaching the first step of the goal. User interface design describes how scripts are generated and executed to create a new warehouse database.

SEMI AUTOMATIC GENERATION OF WAREHOUSE SCHEMA

As stated previously, the goal of this projects it to be able to create and populate databases for data warehouses. This also involves creating a data warehouse schema and loading the warehouse with subsets of relevant data from the source database. The proposed tool to allow users to select, extracts, clean, and converts data from source system structures into consistent target warehouse data structures. In addition, the data from the source database is populated into the target database. The data warehouse can be populated on a frequency that meets the organization's needs. A data warehouse depends totally on its ability draw information from across the organization. The proposed tool provides users with the ability to connect to any source database to draw the required information. Information is drawn into the warehouse by consolidating and cleansing data before populating the warehouse database. This is done automatically after users finalize the target database schema and the mapping with the source database schema.

Data warehousing involves mapping subsets of relevant data from the source database to the target database. The target database schema is designed based on the data that is being transported from the source database. Hence, there is a mapping between the structure of the source database and the target database. This mapping is termed as Schema Mapping. For a given source database, the tool helps users in arriving at an appropriate mapping to create a semantically related warehouse. After a mapping has been formalized, a new warehouse is created. Then, relevant data is automatically transported from the source database to the newly created warehouse. Each mapping created for a source and target database is stored in XML files. This ensures that users can make further changes to the mapping by loading them later.

This tool provides users with a graphical interface to perform Schema Mapping:

- Select source database: Specify the master database for which the warehouse needs to be created,
- Select tables: Users may select only relevant tables from the source database to be a part of the target database. Figure5 presents a mapping diagram.
- Enforce referential integrity: If child tables are selected to be a part of the target database, the corresponding master tables need to be selected too. The master tables are internally computed by the tool and are selected automatically.

CONCLUSIONS AND FUTURE WORK

The SAMDSG thesis tool plays a vital role in providing support for automated data warehouses [1]. It is simple to use, highly interactive and provides an easy means to creating a new data warehouse. It also acts as a reliable tool to quickly explore schema of the source database in order to generate schema for the data warehouse. The SAMDSG tool underlying complex mechanisms from its users, except where it is absolutely appropriate and necessary to expose them. In effect, even non-technical users can create, populate and update data warehouses with minimal time and effort. Attributes from source tables can be mapped into new attributes in the warehouse database tables using aggregate functions. Then, relevant data is automatically transported from the source database to the newly created warehouse. The tool thus integrates warehouse creation, schema mapping and data population into a single general-purpose tool.

This tool has been designed as a component of a framework, whose users are Database Administrators. They will also be able to synchronize updates of multiple copies of the data warehouse. Warehouse images that need to be updated are taken offline and applications that need to access the data warehouse can now access any of the other image warehouses. The Image Switcher built into this framework switches between databases in a way that is very transparent to applications so that they do not realize existence of multiple copies of the data warehouse. It also ensures that ongoing transactions on a particular database are not interrupted when the database is scheduled to be taken offline. The thesis tool is system independent. This gives it advantages such as portability and wide application. This tool can access any database with minimal effort since there is no hard coding of information in the application.

REFERENCES

1. ***LGI systems incorporated: A definition of Data Warehousing***. Retrieved from http://www.dwinfocenter.org/defined.html

2. LGI Systems Incorporated: Maintenance issues for Data Warehousing Systems. Retrieved from http://www.dwinfocenter.org/maintain.html

3. Gittleman, A. (2002). ***Advanced Java Internet Applications*** (2nd ed.). Scott Jones Inc.

4. Hunter, J., & McLaughlin, B. (2000 May). ***Easy Java-XML integration with JDOM.*** Retrieved from http://www.javaworld.com/javaworld/jw-05-2000/jw-0518-jdom.html

5. Chaudhuri, S., & Dayal, U. (1997). An overview of data warehousing and OLAP technology. ***ACM SIGMOD Record*****,** 26, 65-74.

6. Cuzzocrea, A., Sacca, D., & Serafino, P. (2006). A hierarchy driven compression technique for advanced OLAP visualization of MD data cubes. In *Proc. of* ***8th Int'l Conf. on Data Warehousing and Knowledge Discovery*** (DaWak), (pp. 106-119). Springer Verlag.

7. Goil, S., & Choudhary, A. (1997, December). High performance OLAP and data mining on parallel computers. ***Data Mining and Knowledge Discovery*****,** 1(4), 391-417.

8. Peralta, V., Marotta, A., & Ruggia, R. (2003, June). ***Towards the automation of data warehouse design*** (Technical Report TR-03-09). InCo, Universidad de la República, Montevideo, Uruguay.

9. Sarma, Mani V., & Premchand. (2011, November 2). MD Context –Dependent information Delivery on the Web. ***Advanced in Computational Sciences and Technology (ACST),*** ISSN 0973-6107, 4, 131-143.

10. Sarma, Mani V., Premchand., Rao, Purnachandra V., & Goud, Sudheer V. (2011, November 3. Integrating DW with Web based OLAP Operations to develop a Visualization module for hotspot clusters. ***Advanced in Computational Sciences and Technology (ACST),*** ISSN 0973-6107, 4, 341-358.

11. Sarma, Mani V., Rambabu, D., Premchnd, & Rao, Purnachandra V. (2011). A New User Frmework to design MD model from Relational One. ***Journal of Computer Applications(JCA),*** ISSN 0974-1925, 4(IV), 127-130

12. Sarma, Mani V., Rambabu, D., Premchnd, & Rao, Purnachandra V. (2012). A MD model for automatic Schema Generation process by integrating OLAP with Data mining Systems. ***International Journal of Information & Computation Technology***, ISSN 0974-2239, 2(1), 27-36.

*****

BRANDING OF B-SCHOOL'S IN INDIA: A CASE STUDY

Md. Mahmood Ul Farid[27]

ABSTRACT

Education system in India dates back to a time when no fee where charged. Voluntarily "grudakhshina" was prevailing in the society. Today education becomes highly commercialized. As we all know management education were transformed completely in 10 to 15 years.

Earlier there were a few institutes but now with the increase of institutions demand and supply mismatch happened resulted in intense competition. According to Samir Barua (Director IIMA) "Foreign universities is the biggest challenge once they are allow to set up campuses in India".

It is very difficult to differentiate because education as a service cannot be patented. Now business schools are bound to apply the innovative branding strategies, which they teach to their students for different other business activities.

KEYWORDS

Education Marketing, B-School, Marketing, Management Education, Brand, Schools etc.

INTRODUCTION

A brand refers to the "name, term, sign, symbol, or design, or a combination of them, intended to identify the goods and services of one seller or group of sellers and to differentiate them from those of competition" (Keller, 2008, p.7). The components identifying and differentiating brands include brand names, URLs, logos, symbols, characters, spokespersons, slogans, jingles, packages, and signage. These brand elements are managed strategically in many industries in order to convey clear and consistent images and value propositions to consumers, but B-schools have been slow to consider the value of developing strong brands. B-schools should move from a "reputation building" perspective to brand management.

A high-quality faculty and students construct a successful reputation, strong image in the corporate sector, prominent alumni and high survey rankings. Successful brands were perceived to have good reputations. B-schools should benchmark their reputation based on the percentage of full-time staff with PhDs, research rating based on academic publications, entry fees, ratio of full-time staff to full-time students, academic aptitude of students, starting salary of students after graduation, percentage of students employed upon graduating, and the B-school overall reputation.

IMPORTANCE OF BRANDING

People in management education in India fail to realize that in an era of self-financing colleges, on-line education and foreign tie-ups, the creation of brand images for B-School is paramount importance. The degrees of IIMs, IIT's, Kellogg's and MIT are in greater demand is only because of the quality of the experience which the student undergoes in these institutions before getting the same degrees, which are offered elsewhere. Some of the top education brands of the world are in fact richer than some nations.

According to a survey by the National Association of college and university Business offers (NACUBO), "Harvard is right at the top with an endowment estimated at $26 billion bigger than the GDP of Syria, Sri Lanka and Sudan". The universities, which follow that, are Yale (15.2b), Stanford ($12.2 billion), University of Texas (11.6 billion) and Princeton ($11.2 billion). In India, it is time we should work on the Branding issues of the education, which will make our institutions Famous, rich, and most sought after educational hub and will certainly become a Global Brand.

Education can be an important source of national income, provided it is promoted globally like Australia Singapore and USA etc. Global management education market is estimated to be more then USD 40 Billion by 2020.

COMPETITIVE POSITIONING

Success of management education is depend on sustainable business models and selecting and implementing right strategies i.e. right teaching models for course delivery. Institutes must offer different courses for different segments like courses for basic learner's job aspirants, people needing to upgrade their knowledge, customize courses for corporate customers. Such as XLRI is, the only institutes were the study of business ethics is compulsory in all programmes. IFIM Bangalore included health and fitness in their curriculum as part of personality programme. IIMB undergoes a course called business governance and society (BGS). It is also necessary to develop and nurture highly specialize faculty course content should be aligned with real life demand.

[27]Research Scholar, University Department of Commerce and Business Administration, T.M. Bhagalpur University, Bihar, India, mahmood.mba@gmail.com

BRANDING

The concept of branding applied in higher education, is some-what different from branding in the commercial sector. It is about "who we are", and is not limited to what a particular product offers the marketplace. An educational brand is often equated to institute's academic reputation.

The values-centric approach inherent in branding provides an institution with an anchor to guide responses to constituent needs and expectations. The brand is the result of appropriate strategic directions, resource allocations, hiring decisions, and curriculum development. Those institutes serve determines the value of educational brand. Worth of the brand determined by what students are willing to pay for it, donors are willing to give to support it and faculty and staff are willing to contribute to make it real.

There are two major components of branding in Education:

- Promotion of the brand,
- Delivering on the promise of the brand.

PROMOTION OF BRAND

Before the promotion, the desired brand identity must be defined. A brand rationale, brand attributes, and brand benefits should be clearly articulated and consistently reflect the institution's values while aligning with constituent expectations. The brand tactics are aligned with the brand strategy and that brand strategy is aligned with the institution's mission and values.

Five things we can do for Promotion:

1. Understanding the need of customer: Surveys, focus groups, observations, a review of historical data can be used to understand their needs.

2. Identifying the most valued market segments: Defining the characteristics of each segment, including motivators and barriers supporting the institution's objectives.

3. Identifying the weak links of 'the brand': To exploit the most out of target segment, brand should be modified according to the demand.

4. Identifying the strength of brand and effectively position it against the competitors. Identify and use institutional strengths and competitor weaknesses in particular market segment.

5. Differentiating the institution through relevant communications. Describe how their unique needs will be met by your institution's value proposition and how it is different from competitors.

Beyond this proper communication, channels must be used for most effectively delivering the message. Most appropriate time, means and the infrastructure should be decided for the promotion. Quality of execution should be monitored and effectiveness of the brand promotional efforts should be measured.

Even with careful planning and near-perfect execution, brand promotion is failed if a prospective students experience with the institutes is incongruent with the brand message. Delivering on the promise of the brand is the single most important aspect of branding a B-school.

DELIVERING ON PROMISE

Many B-school marketing professionals fail to understand their brand. Ultimately, fail to manage their brand. Every institute has a brand, but it suffers from benign neglect. To effectively shape how constituents view an institution, it must begin first by understanding the promise inherent in the existing brand or the brand the school aspires to have. There must be a similarity between what an institution claims to be and what its constituents actually experience in the campus.

After understanding the 'brand promise', these steps should be taken to ensure consistent delivery:

1. **Defining the Brand Promise**

   It must be based on the institution's personality and core values. Brand promise must be relevant both to internal and external constituents. Employees must passionately believe in and care about the promise for it and delivered through the educational experience and student services. The promise must be influential enough to be accepted and practiced by different individuals with their own unique beliefs and values. In the academy, this is the only means to have a balance between the objective of universal adoption and maintaining a modest autonomy.

2. Living the Brand Promise

All faculties, staffs, and administrators are the "brand ambassadors" of Institutes." Their interaction with students and outer world either can foster or erodes institutional trust. Since every valued relationship is built on a foundation of mutual trust. Therefore, the students, their families, the school's alumni, and others we serve must have trust on the institute.

3. Brand Promise Should Be Operationalized

The promise must be personified through the services, business transactions, human interactions, information delivery, and learning experiences. It must be in the culture and become a part of the institution's DNA.

4. **Consistency in delivering Brand Promise**

To achieve it institutions must:

- Clearly define the desired constituent experience,
- Ensure the employee experience is aligned with the desired constituent experience,
- Create conducive environment for employees to feel passionate about the organization and its promise. Improve the campus environment to values the contributions of individuals, which proactively enhance human capacity.

5. Brand Promise Should Be Conveyed

For effectively conveying of promise requires an ongoing internal and external campaign. It requires careful management of constituent expectations.

CONCLUSIONS

Marketing department can implement only the promotional aspect of branding. Brand promotion without the brand promise is often counterproductive. When the expectations generated from promotional activities are incongruent with constituent reality, the image of the institution is tarnished and trust is eroded. Branding is about systemic institutional change. Branding should be used as a catalyst for defining who the institution is and what it aspires to become. Branding can be a means of unifying the campus around a common purpose and vision. Branding effort must cast in such a context that it became more palatable, such as "institutional promise" or "constituent engagement." With acceptable terminology, holistic approach, and the necessary previous principles for success radically improve an institution's image and future reality.

REFERENCES

1. (2010). ***AACSB International Issues in Management Education Survey 2010.*** Internal Publication, Graduate Management Admission Council. Retrieved from www.gmac.com

2. Christine, Wright-Isak. (2010, June 23). ***The Intersection of Scholarship and Practical Experience.*** Marketing Florida Gulf Coast University.

3. (2003, November/December). The zen of B-Branding, Tricia Bisoux, ***BizEd.***

4. Association of American Colleges and Universities (2007). ***National Leadership Council for Liberal Education and America's Promise.*** College Learning for the New Global Century.

5. Gibbs, N. & Thornburgh, N. (2006, August 13). Who needs Harvard? ***TIME***.

6. (2001). ***Returning to Our Roots: Executive Summaries of the Reports of the Kellogg Commission on the Future of State and Land-Grant Universities.*** Kellogg Commission on the Future of State and Land-Grant Universities.

*****

FOR ANY CLARIFICATION OR SUGGESTION, WRITE US:

Editor-In-Chief
Pezzottaite Journals,
24, Saraswati Lane, Bohri, Near Modern Dewan Beverages,
Jammu Tawi – 180002, Jammu and Kashmir, India.
(Mobile): +91-09419216270 – 71
editorinchief@pezzottaitejournals.net
contactus@pezzottaitejournals.net

A SURVEY OF DEEP FRAGMENTATION TECHNIQUES FOR TEXTURED IMAGES

K. Meenakshi Sundaram[1] C. S. Ravichandran[2]

ABSTRACT

In this paper, it is intended to summarize and discuss the methods of fragmentation for textured images in various applications of image processing. In particular, the problem of real time approaches for textured images is analyzed and their performances are studied and discussed. The main aim of fragmentation is to locate objects of interest based on the available criteria and it is sometimes a computer vision problem, but many important segmentation algorithms are too simple to solve this problem accurately. They compensate for this limitation with their predictability, generality, and efficiency.

KEYWORDS

Fragmentation, Segmentation, Textured Images etc.

I. INTRODUCTION

DIGITAL image processing is enhancement by improving image quality by filtering etc.; and restoration by compressing data to save storage and channel capacity during transmission. It is a rapidly evolving field with growing applications in science and engineering. It holds the possibility of developing the ultimate machines that could perform the visual functions of all. Image analysis technique, the same input gives out somewhat detail description of the scene whose image is being considered. Most of the image analysis algorithms perform segmentation as a first step towards producing the description. The segmentation of textured images is a long standing problem in Computer Vision, which has been addressed from various perspectives, with variational models, and MRFs, [1] being the most common approaches. Although texture is a fundamental characteristic of a region, the complexity involved in its quantification has prevented its effective incorporation into the segmentation process. Segmentation divides the special domain, on which the image is defined, into meaningful parts or regions. The level to which this subdivision is carried depends on the problem solved i.e. segmentation should stop when the objects of interest in an application have been isolated. Generally segmentation is one of the most difficult tasks in image processing. There exists no general segmentation algorithm which can work reasonably well for all images. The suitable segmentation algorithm for the particular problem must be chosen or developed as they are ad hoc in nature. The algorithms may be incorporated explicitly or implicitly, or even in the form of various parameters. These algorithms are based on satisfying homogeneity property or detecting abrupt changes in image features or both approaches. This step in the process of analysis determines the eventual success or failure. Generally these techniques can be classified based on three different aspects: First the model used for segmenting (explicitly stated), secondly, the optimized criterion (implicitly stated) and finally, the algorithm which is employed to compute the segmentation. A large variety of segmentation methods can be found in the literature. Of more interest to us are the segmentation methods which try to retrieve regions using different models [2]. The idea of using a combination of different segmentations to obtain the best segmentation of an image has been suggested by Cho and Meer [3]. However, they make use of small differences resulting from random processes in construction of a Region Adjacency Graph (RAG) pyramid to generate their segmentations.

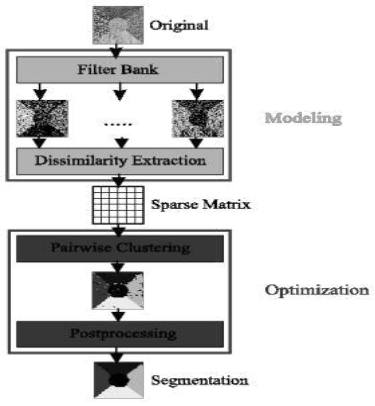
Algorithms for subsequent image processing stages like motion analysis and tracking, stereo vision, object recognition and scene interpretation often rely on high quality image segmentation [4]. For textured images one of the main conceptual difficulties is the definition of a homogeneity measure in mathematical terms [5]. A number of distinct approaches have been suggested for textures images falling into two major classes [6]. We have four traditional methods for image segmentation, namely, Fuzzy C-Means clustering based segmentation, Region split and merge segmentation, Region growing segmentation and Histogram thresholding based segmentation. Some other methods are Pixel based Segmentation; Model based Segmentation, Multi-scale Segmentation, and Semi-automatic Segmentation etc.

II. OBJECTIVES OF STUDY

The objective of this paper is to review the methodologies of segmentation of real time approaches for textured images. The aim of segmentation is to locate objects of interest based on the available criteria and it is sometimes considered as a computer vision problem. Rest of paper is described as follows: Section 3rd discusses the segmentation process. The review of previous work is discussed in Section 4th. Section 5th discusses the summary of methods and Section 6th discusses the review evaluation. Conclusions are drawn in last Section.

III. FRAGMENTATION PROCESS

Figure-1: Fragmentation



Sources: Authors Compilation.

Fragmentation is based on four cascaded design decisions, concerning the image representation, texture homogeneity, objective functions and optimization procedures. Figure-1

[1] Associate Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, meenaksji@gmail.com
[2]Professor and Dean, Sri Ramakrishna Engineering College, Coimbatore India, dean-eee@srec.ac.in

gives the process involved in segmentation. The original image forms the input, and modeling is performed by two steps such as reducing the noise, and dissimilarity extraction based on texture homogeneity. The image is represented as sparse matrix. Based on the objective function and optimization procedures, pair wise clustering is performed and after post processing steps, the segmented image would be available for analysis

IV. RELATED WORKS

Texture segmentation is an important task in many computer vision applications. Texture can be described in a variety of ways. It can be generated by primitives organized by placement rules, or as the result of some random process. It can be found in a continuous spectrum, from purely deterministic to purely stochastic. Texture is an important surface characteristic and based on it shape and motion can be estimated. Taxonomy of problems encountered within the context of texture analysis could be that of classification / discrimination, description, and segmentation. These problems are listed in order of increasing difficulty, and it is clear that the major problem is that of texture segmentation. Generic segmentation is the partitioning of an image into regions that are "homogeneous" with respect to one or more characteristics [7]. The process of segmentation is rigorously defined by Pavlidis and its corresponding computational complexity is discussed by Gurari and Wechsler. A basic issue to be considered is that of the cell unit size, i.e., the resolutions of the area over which measurements are taken in order to test for homogeneity.

The segmentation is achieved in two phases: the first one consists of evaluating, from disjoint blocks which are classified as homogeneous, the model parameters for each texture present in the image. This unsupervised learning phase uses a fuzzy clustering procedure, applied to the features extracted from every pixel block, to determine the number of textures in the image and to roughly locate the corresponding regions. The second phase consists of the fine segmentation of the image, using Bayesian local decisions based on the previously obtained model parameters. The problem considered in Unsupervised Segmentation of textured images by Edge Detection in Multidimensional feature operating in unsupervised mode. A set of m features are used to represent such characteristics. Consequently, the problem is transformed into edge detection problem in multidimensional feature space.

In addressing the problem, it requires resolution of two major issues: namely, 1) selection of 'm' features to represent texture, and 2) selection of an appropriate local size for texture measurement. The segmentation algorithm which is presented here can employ any set of m features provided that they possess the following three properties: 1) strong within-class invariance, 2) strong between - class separation, and 3) low sensitivity to small sample size, meaning that: satisfy requirements 1) and 2) should not be contingent upon availability of large size texture fields. The selected features are the estimated parameters of a class of spatial interaction non causal random field models called the "Simultaneous Autoregressive (SAR)" model which is fitted to the local region under consideration. A second and very important problem is how to select an appropriate window size over which local textural characteristics are observed and measured. The window should be large enough to contain several texture elements for the enclosed region to exhibit similar textural characteristics as that of the underlying region that it masks. At the same time, it should be as small as possible to enable accurate detection of the edges. Operating in an unsupervised mode prevents a priori selection of a good window size. In fact, a major shortcoming of the previously developed edge-based algorithms for texture segmentation is the arbitrary pre selection of this parameter. A systematic method is used in SAR model which enables the algorithm to automatically select an appropriate window size. Other approaches besides edge-based ones include studies relying on estimation theory, clustering in the feature space and the split-and-merge technique.

Segmentation is approached as a statistical approach problem [8]. These methods of segmentation use MRF to model the discrete field containing the individual pixel classification. The result is quite difficult to maximize globally. For global optimum Stochastic relaxation algorithms have been applied. Through a reasonable method between performance and computation, greedy minimization algorithm is known as estimation by Iterated Conditional Modes (ICM) is efficient. The Multiple Resolution Segmentation (MRS) is less likely to be trapped in local minima than ICM and also requires less computation. The above mentioned methods are good than Multi grid methods. For individual texture GAR model is used for segmentation as well as parameter estimation and for multiple resolution techniques extract texture features for each pixel average it to yield an aggregate statistic for the region and is well adapted to use with it.

Todd and Harry addressed the generic issue of clustering / grouping. Recent research, both in computer and human vision, suggests the use of joint spatial / spatial-frequency (s/sf) representations. The spectrogram, the difference of Gaussians representation, the Gabor representation, and the Wigner distribution are discussed and compared. It is noted that the Wigner distribution gives superior joint resolution proving the feasibility of using s/sf representations for low-level (early, pre-attentive) vision.

Iasonas [9] incorporates recent results from AM-FM models for texture analysis into the variational model of image segmentation and examines the potential benefits of using the combination of these two approaches for texture segmentation. Using the Dominant Components Analysis (DCA) technique they obtained a low-dimensional, yet rich texture feature vector that proves to be useful for texture segmentation. By using an unsupervised scheme for texture segmentation, where only the number of regions is known apriori, synthetic and challenging real-world images results demonstrate the potential of the proposed combination. The algorithm for segmentation of textured images using a Multi-resolution Bayesian approach uses a Multi-resolution Gaussian Autoregressive (MGAR) model for the pyramid representation of the observed image, and assumes a multiscale Markov random field model for the class label pyramid. In unsupervised segmentation strategy for textured images, based on a hierarchical model in terms of discrete Markov Random Fields, the textures are modeled as Gaussian Gibbs Fields, while the image partition is modeled as a Markov Mesh Random Field [10].

In study of segmentation both complex and real filters were used [11]. Complex prolate spheroidal sequences were used as channel filters, and channel envelopes were extracted to form a feature vector. In another study complex Gabor filters were

applied and envelope and phase information were extracted from two quadratic components of distinct output channels. Use of real Gabor filters was reported, where feature extraction included a nonlinear transformation, $r > (f) = \tanh(at)$, and a statistical measurement of average absolute derivation was performed on overlapping windows. However, even this sophisticated approach of feature extraction had limitations. Although Gabor filters possess desirable properties for this application, recent developments in wavelet theory provide an alternative approach with several advantages as listed below:

- Wavelet filters cover *exactly* the frequency domain (provide a mathematically complete representation).
- Correlations between features extracted from distinct filter banks can be greatly reduced by selecting appropriate filters.
- Adaptive pruning of a decomposition tree makes possible the reduction of computational complexity and the length of feature vectors.
- Fast algorithms are readily available to facilitate implementation.

Gabor filters adopts real wavelet packet frames (tree structured filter banks) for channel filters, and introduces two envelope detection algorithms for feature extraction. By using color image in textured analysis [12] is to incorporate the chromatic information into texture analysis, assuming that the RGB color space is used, the following choices exist:

- Each color band (i.e., R; G; B) is processed separately.
- Information across different bands (e.g., cross-correlations RG, RB, GB) is extracted.
- Both individual color band and cross-band information is used.
- A composite measure to describe the chromatic information is used.

The fourth alternative is explored using the xyY color space. The main goal of the system is to separate a given image into two parts, namely, a Region of Interest (ROI), and the rest of the image (i.e., the background. The system performs analysis on luminance and chrominance in parallel, and, at the final stage, results are combined to detect changes (i.e., loss/gain) in a specific area of the image (ROI). Processing starts by transforming a given image from RGB to xyY. This produces the luminance component (Y) directly, whereas the two chromaticity values (x; y) are combined to provide for single-valued chrominance. Textural information, such as sizes and orientations of basic image features (e.g., edges, blobs), is contained in the luminance component. Thus, a set of filters tuned to different sizes and orientations is applied on luminance and produces a corresponding set of filtered images. Smoothing of the filtered images follows, thus, eliminating spurious / negligible regions. The smoothed images are combined into a single image, based on a neighborhood pixel similarity measure, and boundaries of potential ROI's are extracted using a perceptron-type processing mechanism. The result of luminance processing is, thus, a Boundary Image. Chrominance processing proceeds in two stages. First, the chrominance histogram is computed and multiple thresholds are identified.

Secondly, these thresholds are used to segment chrominance image into a corresponding number of regions (i.e., potential ROI's). Thus, the result of chrominance processing is a Region Image. Using a region expansion algorithm, the Boundary and Region Images are combined to locate the desired Region of Interest. The result is a ROI Image showing the identified ROI. The final stage involves the comparison of two or more ROI images to locate possible scene changes. Typically, two or more images of the same real-world scene are taken at different times. Each of these images will result in a corresponding ROI image, after going through the various segmentation stages (i.e., luminance and chrominance processing). Change detection and measurement is performed by comparing two such ROI Images using logical pixel operators. The end result is threefold:

- Incorporation of texture and color attributes for scene analysis;
- Development of computationally efficient and easily implementable algorithms for analysis of color textures;
- Development of appropriate neural network architectures for image segmentation & classification.

This methodology provides the analysis component of an autonomous system. It incorporates color and texture visual attributes into a unified framework and utilizes them to detect and measure loss/gain.

Recently, the multichannel / multi-resolution approach has drawn a lot of attention. There is evidence that images are decomposed into a collection of band pass sub-images by simple visual cortical cells to form features for segmentation task. Gabor filters are suitable for such decomposition because their impulse responses and the joint space / spatial-frequency resolution are optimal. Though successful results using a large set of Gabor filters that cover a half frequency plane with the orientation and frequency selectivity requirements have been reported, the computational effort and storage requirement cause major problems, furthermore, if the filters' parameters do not match the spectral characteristics, the segmentation results are usually unsatisfactory. To eliminate these shortcomings, it is desirable to reduce the number of Gabor filters, which may not cover a half frequency plane in full yet still capture significant spectral information by tuning the parameters adaptively. Many tuning algorithms have been proposed [13] through a global Fourier analysis or a spectral feature contrast analysis. More recently, wavelet theory has been developed enabling a new space / spatial-frequency analysis. Mallat applied the wavelet transform with an efficient pyramid structured algorithm to texture analysis in a multi-resolution framework. Unser proposed an over complete wavelet transform by incorporating redundant information.

V. REVIEW EVALUATION

Here, the gray level co occurrence matrix (GLCM) type of features is considered rather than SAR features. It has been very popular in texture analysis. Six such features, namely, energy, contrast, entropy, correlation, homogeneity, and cluster shade are utilized. The performance of GLCM algorithm is quite satisfactory. However, it should be pointed out that since a window-based technique is utilized, one cannot expect the algorithm to locate edges with pixel level accuracy. Errors on the order of several pixels should be tolerated. It takes around 25 min to segment an image on a VAX 111750. However, the extensive computation time is not really a drawback since the algorithm is very easily implementable on a parallel processor.

The MRS algorithm is better than ICM as it requires less computation though the algorithm performs comparably to Simulated Annealing. MRS yields improvement when the information in pixel is low, else it must be combined over larger regions to correctly segment the images. Usually, the MRF contains the discrete class of each pixel in the image. The objective then becomes to estimate the unknown MRF from the available data. In practice, the MRF model typically encourages the formation of large uniformly classified regions. Generally, this smoothing of the segmentation increases segmentation accuracy, but it can also smear important details of segmentation and distort segmentation boundaries. Approaches based on MRFs also tend to suffer from high computational complexity. The non causal dependence structure of MRFs usually results in iterative segmentation algorithms, and can make parameter estimation difficult. Moreover, since the true segmentation is not available, parameter estimation must be done using an incomplete data method such as the EM algorithm.

A tradeoff exists between over segmentation, partition into too many regions, and under segmentation, in which case larger regions are obtained at the expense of possible erroneous fusions. A complete segmentation system is the complex and makes use of many heuristics. To reduce over segmentation, in the absence of context dependent information, probabilistic models are required to guide the fusion process. The difficulty of segmentation is an aspect of the local / global duality problem. A region is declared homogeneous by analyzing small local neighborhoods. The larger these neighborhoods, the more reliable are the extracted spatial statistics *given* that the data in the neighborhood is indeed homogeneous. On the other hand, using a larger neighborhood increases the chances of analyzing non homogeneous data under the assumption of homogeneity.

While providing a framework for predicting visual grouping effects, the Gestalt laws are not yet grounded in any specific theory of vision. Wigner distribution performs the superior joint resolution. The s/sf representations, in particular the WD, seem to be appropriate for performing such tasks. An experimental system based on the pseudo-Wigner distribution was implemented, and experiments in texture segmentation and Gestalt grouping were performed. Evidence was given linking energy content of the primary frequency plane in the PWD with the ability of human subjects to spontaneously discriminate between texture fields. Groupings of elements consistent with the psychophysical predictions of Gestalt laws were achieved. A correlation between the perceived grouping of elements and the energy of associated PWD frequency planes was also demonstrated. This correlation indicates a direction for future research on the relationships between perceptual parameters, such as distance, and perceived groupings. Andrew Laine and Jian Fan tested their representation using an ISODATA clustering algorithm. The number of distinct classes in each textured image was a required parameter for the program. Their test images included samples of two distinct families of textures as Natural textures and Synthetic textures. For the difficult test image, the algorithm achieved outstanding performance. The performance is consistent with the difficulty of segmentation perceived by human observers. It is observed that boundary errors were dependent on shape, i.e., complex boundaries yielded more variance.

George Paschos and Kimon P. Valavanis presented a visual monitoring system that incorporates color and texture processing principles for image analysis. Here the gray-scale information has been given major consideration than chromatic information. Emphasis has been given to its segmentation capabilities which are directly applicable to environments where detection and measurement of change in the sensed scenes is of primary importance. The approach is part of a complete color texture analysis system that includes the described segmentation subsystem as well as additional classification algorithms that form the corresponding color texture classification subsystem. The system can be applied in areas as: 1) automated underwater surveillance in which the visual monitoring system becomes part of a sensor based control architecture of an AUV (Autonomous Underwater Vehicle); 2) wetlands monitoring; and 3) GIS (Geographical Information Systems).

The wavelet transform has many properties such as multi-resolution analysis, fast algorithms, perfect reconstruction etc. that are beneficial for image applications. However, diagonal high pass filter involved in the 2-D wavelet transform gives a strong response to textures with orientations at or close to both + or − 45 degree, which causes ambiguity in textures with symmetric orientations. Motivated by AM-FM representation, the orientation information can be taken into account while decomposing the amplitude function into wavelets (wavelet packets). The results in modulated wavelet (packet) transform, not only generalizes the wavelet (packet) transform but also preserves the desirable properties. The modulated wavelet transform zooms in frequency region centered at modulating frequencies for further decompositions, so texture segmentation can be improved by adapting modulating frequencies to the spectral energy contrast of given image.

Chart-1: Comparison Chart of Assorted Fragmentation Techniques

S.No.	Fragmentation Analysis	Features	Description
1	Multiscale Bayesian Segmentation using a Trainable Context Model	Distort segmentation boundaries, high computational complexity, difficult in parameter estimation	EM algorithm
2	Image segmentation from consensus information	Erroneous fusions, use of many heuristics, over segmentation or under segmentation, duality problem, homogeneity problem.	Bootstrap
3	Segmentation of Textured Images and Gestalt Organization Using Spatial / Spatial-Frequency Representations	The Gestalt laws are not yet grounded in any specific theory of vision.	Wigner distribution
4	Modulation-feature based textured image segmentation using curve Evolution	Image intensity is a poor cue, spurious edges,	Combines the best of the DCA and curve evolution methods

5	Unsupervised Segmentation of textured images by Edge Detection Multidimensional feature	Selection of m features(energy, contrast, entropy, correlation, homogeneity, and cluster shade) to represent texture, and selection of an appropriate local size for texture measurement	SAR model & GLCM algorithm
6	Multiple Resolution Segmentation of textured images	Difficult to maximize globally Performance and computation Tapped in local minima, computation and Good than Multi grid methods Parameter estimation and Multiple resolution techniques	Stochastic relaxation algorithms Greedy minimization algorithm - Iterated Conditional Modes (ICM) MRS GAR model
7	Segmentation of Textured Images using a Multi-resolution Gaussian Autoregressive Model	Pyramid representation Class label pyramid Hierarchical model in terms of discrete Markov Random Fields Image partition	MGAR Multiscale Markov random field model Gaussian Gibbs Fields, Markov Mesh Random Field
8	Frame Representations for Texture Segmentation	Natural textures and Synthetic textures difficult test image	ISODATA clustering algorithm
9	A Color Texture Based Visual Monitoring System For Automated Surveillance	Analysis component of an autonomous system	Incorporates color & texture visual attribute into a unified framework and utilizes them to detect and measure loss / gain
10	Texture Segmentation using Modulated Wavelet Transform	Multichannel / multi-resolution approach fast algorithms, perfect reconstruction	Gabor filters Wavelet packets

Sources: Authors Compilation.

VI. CONCLUSIONS

In this paper, the survey of various fragmentation techniques was discussed for textured images. In computer vision, segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and / or change the representation of an image into something that is more meaningful and easier to analyze Image segmentation is typically used to locate objects and boundaries in images. Future research work is on how to handle the limitations in the algorithm and improve the results.

VII. REFERENCES

[1] Cheng, Hui, and Charles, A., (2001). "Bouman, Multiscale Bayesian Segmentation Using a Trainable Context Model", IEEE Transactions on Image Processing, Volume: 10, Number: 4, April, pp. 511–525.

[2] Hanbury, Allan; Jocelyn Marchadier, and Walter G. Kropatsch, The Redundancy Pyramid and its Application to Image Segmentation, Pattern Recognition and Image Processing Group, Vienna University of Technology, Favoritenstra_e 9/1832, A-1040 Vienna, Austria.

[3] K., Cho, and Meer, P., (1997). "Image Segmentation from Consensus information", Computer Vision and Image Understanding, 68(1): 72-89.

[4] Cho, Kyujin, and Peter Meer, (1997). "Image Segmentation from Consensus Information", Computer Vision and Image Understanding, Volume: 68, Number: 1, October, Article Number. IV970546, pp. 72–89.

[5] Zhang*, Yongyue; Brady, Michael, and Smith, Stephen, (2001). "Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the Expectation - Maximization Algorithm", IEEE Transactions on Medical Imaging, Volume: 20, Number 1, January, pp. 45 – 57.

[6] Corsco, Davide; Fioravanti, Roberto, and Stefan0 Fioravanti, "Morphological Analysis of Textured Images for Identification of Thin structures", DIBE - University of Genova, Italy via Opera Pia 11/A, Genova 1-16145, ITALY.

[7] Todd, R.; Reed and Harry Wechsler, (1990). "Segmentation of Textual Images and Gestalt Organization using Spacial / Spacial-Frequency Representatives", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 12, Number: 1, January.

[8] Bouman, Charles, and Bede, Liu, (1991). "Multiple Resolution Segmentation of Textured Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 13, Number: 2, January, pp 99-113.

[9] Iasonas, Kokkinos, Evangelopoulos, and Petros Maragos, (2004). "Modulation – Feature based Textured Image Segmentation Using Curve Evolution", Proceedings of International Conference Image Processing (ICP-2004) Singapore, 24-27 October, pp. 1201-1204.

[10] Mary, L. Comer, and Edward J. Delp, "Segmentation of Textured Images using a Multi-resolution Gaussian Autoregressive Model", Computer Vision and Image Processing Laboratory, School of Electrical Engineering, Purdue University, West Lafayette, Indiana.

[11] Laine, Andrew, and Jian Fan, (1996). Frame Representations for Texture Segmentation, IEEE Transactions on Image Processing, Volume: 5, Number: 5, May, pp 771-780.

[12] Paschos, George, and Kimon, P. Valavanis, (1999). "A Color Texture Based Visual Monitoring System for Automated Surveillance", IEEE Transactions on Systems, Man, and Cybernetics—Part c: Applications and Reviews, Volume: 29, Number: 1, February, pp 298-307.

[13] Hsi-Chin Hsin, (2000) "Texture Segmentation Using Modulated Wavelet Transform", IEEE Transactions on Image Processing, Volume: 9, Number: 7, July, pp 1299–1302.

*****

A SURVEY ON MONITORING PERFORMANCE IN GRID

Abika S.[3]

ABSTRACT

To process high performance distributed data access and computing in Grid environment, monitoring of resource and network performance is needed. The Grid resources are needed to be discovered, selected and invoked quickly and efficiently in order to satisfy the needs of a demanding environment. This paper discuss how the grid services discover the resources and network performance in a wide area distributed grid and how efficiently complete the submitted job within minimum time. Mobile agents are migrated to start the sensors to measure the network metrics in all Grid Resources from the Resource Broker.

KEYWORDS

Grid Computing, Resources Broker, Network Metrics, Grid Resources etc.

I. INTRODUCTION

A Grid is a collection of computing elements which shares the resources among virtual organizations. As the computing resources, there is a need of building the monitoring system to achieve an effective utilization of the Grid resources. The computational Grids facilitate the software applications to integrate instruments, displays, and computational resources which are managed by various organizations from extensive locations. Network monitoring is very much needed in Grids to avoid the problems which arise due to overloaded servers, failure of network connections, etc.

Network monitoring is the process of monitoring the network metrics such as bandwidth, latency, packet loss rate, throughput, jitter and Round Trip Time (RTT). The network performance monitoring and prediction provides the necessary information for the enrichment of scheduling the best resources such as where to get or put the data and where to execute the job, fault detection and trouble-shooting, identifying the bottleneck, performance analysis and tuning. Processing all monitoring tools for different resources involved with an application, collecting such data, filtering them for obtaining useful information may become a major problem.

II. CLASSIFICATION OF RESOURCE MONITORING MECHANISMS

There are plenty of researches has been done and plenty of method's are available in the field of resource monitoring in grid computing. This survey concentrates on Network monitoring performance of resources in the grid systems. This Survey classifies the monitoring services of a resource and Network performance. Classifications of monitoring service are 6 different methods. They are as follows:

- Relational Grid Monitoring Architecture (RGMA),
- Net Logger,
- Java Agent For Monitoring and Management (JAMM),
- Simple Network Management Network Protocol (SNMP),
- Data Intensive And Network Aware (DIANA),
- Inter-Domain Meta-Scheduling in Grid.

Relational Grid Monitoring Architecture (RGMA)

The Relational Grid Monitoring Architecture (R-GMA) [1] as part of the Data Grid project, to provide a flexible information and monitoring service for use by other middleware components and applications. R-GMA presents users with a virtual database and mediates queries posed at the database users pose queries against a global schema [2] and R-GMA takes responsibility for locating relevant sources and returning an answer. R-GMA's architecture and mechanisms are general and can be used wherever there is a need for publishing and querying information in a distributed environment. The advantage of R-GMA's approach is that users are offered all the flexibility that the relational model and SQL query language bring. The relational approach has a sound theoretical basis, and its frame work has been extended recently to the world of streams by the database community. In R-GMA [6], both "one-time" and continuous queries over streams are supported. The flexibility gained by choosing a relational data model over, say, a hierarchical model does come with some costs. It is less obvious how a relational schema can be distributed across the Grid. Indexes still need to be chosen with certain queries in mind [3]. However, the main advantage is that the relational model allows queries to explore complex relationships in data.

The disadvantage of software to produce the resource information was developed within data Grid & generated data in LDAP format. The concern about the performance and scalability of MDS, it was decided to migrate to R-GMA. The R-GMA create redundancy of data since the LDAP database were essentially identical. The performance of R-GMA & stability was hardening the code.

Net Logger

Lawrence Berkeley National Lab developed the Net Logger Toolkit [4], which is designed to facilitate non-intrusive instrumentation of distributed computing components. Using Net Logger, distributed application components are modified to produce time stamped logs of "interesting" events at all the critical points of the distributed system. Events from each component are correlated, which allows one to characterize the performance of all aspects of the system and network.

Net Logger is used to instrument Grid applications and services, and includes the ability to change the logging-level on the fly by periodically examining a configuration file. The Net Logger binary data format provides an extremely efficient, light-weight transport mechanism for the monitoring data. Py GMA provides an easy to use, SOAP-based framework for control messages. Py GMA also provides a standard publish-subscribe API for Grid monitoring event publication. The Net Logger instrumentation library is very efficient and easy to use. Using the binary format, Net Logger can serialize on the order of half a million events per second. In order to instrument an application to produce event logs, the application developer inserts calls to the Net Logger API at all

[3] PG Scholar, Karunya University, Tamil Nadu, India, abikasamuel@gmail.com

the critical points in the code, then links the application with the Net Logger library. This facility is currently available in several languages: Java, C, C++, Python, and Perl. The API has been kept as simple as possible, while still providing automatic time stamping of events and logging to either memory, a local file, syslog, a remote host. The ability to monitor & manage distributed computing component is critical for enabling high performance distributed computing system.

Java Agent for Monitoring and Management

The automated agent-based architecture that has developed is called Java Agents for Monitoring and Management (JAMM) [5]. The agents, whose implementation is based on Java and Java Remote Method Invocation (RMI), can be used to launch a wide range of system and network monitoring tools, and then to extract, summarize, and publish the results. The JAMM system is designed to facilitate the execution of monitoring programs, such as net stat, io stat, and vm stat, by triggering or adapting their execution based on actual client usage. On-demand monitoring reduces the total amount of data collected, which in turn simplifies data management. For example, an ftp client connecting to an ftp server could automatically trigger net stat and vm stat monitoring on both the client and server for the duration of the ftp connection. Application activity is detected by a port monitor agent running on the client and server hosts, which monitor traffic on a configurable set of ports, use the JAMM system to generate monitoring data for analysis by Net Logger, a toolkit for performance analysis of distributed systems. When performing a Net Logger analysis of a complex distributed system such as a Computational Grid, activating all desired host and network monitoring can be a very tedious task.

Fault detection and recovery mechanisms need monitoring data to determine [6] if a server is down, and whether to restart server or redirect service requests elsewhere. A performance prediction service might use monitoring data as inputs for a prediction model, which would in turn be used by a scheduler to determine which resources to use. As distributed systems such as Computational Grids become bigger, more complex, and more widely distributed, it becomes important that this monitoring and management be automated. The problem is to use a collection of software agents as an event management system which is designed for use in a Grid environment. The securely start monitoring programs on any host, and manage their resulting data. It can provide a standard interface to host monitoring sensors for CPU load, interrupt rate, TCP retransmissions, TCP window size, and so on. Agents can also independently perform various administrative tasks, such as restarting servers or monitoring processes.

In JAMM model, one can add additional event gateways, and additional sensor directories as needed, reducing the load where necessary. In the case where many consumers are requesting the same event data, the use of an event gateway reduces the amount of work on and the amount of network traffic from the host being monitored. An event gateway would typically be run on a separate host from the grid resources, to ensure that the load from the gateway did not affect what was being monitored. JAMM creates a number of security vulnerabilities which must be analyzed and addressed before such a system can be safely deployed on a production Grid.

The users of such a system are likely to be remote from the machines being monitored and to belong to different organizations [7]. Users want to find out what sensors are running and how to subscribe to their event data; users may need to cause sensor programs to be started or to generate a higher level of data collection and finally users want to subscribe to sensor data via an event gateway. In each case the domain that is being monitored is likely to want to control which users may perform which actions. The main drawback of JAMM is not support the replication.

Simple Network Management Protocol (SNMP)

The three types of network measurement and monitoring are active measurement, passive measurement, and SNMP-based measurement. SNMP [8] is a popular protocol for network management. It is used for collecting information from, and configuring, network devices, such as servers, printers, hubs, switches, and routers on an Internet Protocol (IP) network. SNMP can collect information such as a server's CPU level. The SNMP protocol was designed to provide a "simple" method of centralizing the management of TCP/IP-based networks. To manage devices from a central location, the SNMP protocol is facilitates the transfer of data from the client portion of the equation to the server portion where the data is centralized in logs for centralized viewing and analysis

Data Intensive and Network Aware (DIANA)

In Grids scheduling decisions are often made on the basis of jobs being either data or computation intensive: in data intensive situations jobs may be pushed to the data and in computation intensive situations data may be pulled to the jobs [9]. This kind of scheduling, in which there is no consideration of network characteristics, can lead to performance degradation in a Grid environment and may result in large processing queues and job execution delays due to site overloads. Data Intensive and Network Aware (DIANA) [10] meta-scheduling approach, which takes into account data, processing power and network characteristics when making scheduling decisions across multiple sites. The DIANA meta-scheduler keeps track of the load on the sites and selects a site which has a minimum load and queue and has the desired data, processing capability and network stability. It can be difficult to estimate the true effect of the DIANA scheduling approach if jobs are run at different times, and the results of various approaches are taken at different times.

DIANA scheduling is equally applicable to short and long duration jobs. For longer jobs it is the execution time which will vary and accordingly queue times will also increase. The job scheduler provides a global ranking of the computing resources and then selects an optimal one on the basis of this overall access and execution cost.

The DIANA approach considers the Grid as a combination of active network elements and takes network characteristics as a first class criterion in the scheduling decision matrix along with computations and data. If data are not replicated to the site where the job is supposed to be executed, the data need to be fetched from remote sites. The main demerits of the DIANA is once the job gets a CPU, then abort that job and move that job to other Grid site is difficult because it uses non pre-emptive mode of execution.

Inter-domain Meta-scheduling in Grid

Grid computing generally involves the aggregation of geographically distributed resources in the context of a particular application. The resources can exist within different administrative domains requirements on the communication network must also be taken into account when performing meta-scheduling, migration or monitoring of jobs [16].

Similarly, coordinating efficient interaction between different domains should also be considered when performing such meta-scheduling of jobs. A strategy to perform peer-to-peer-inspired meta-scheduling in Grids is presented. This strategy has three main goals: (1) it takes the network characteristics into account when performing meta-scheduling; (2) communication and query referral between domains is considered, so that efficient meta-scheduling can be performed; and (3) the strategy demonstrates scalability, making it suitable for many scientific applications that require resources on a large scale.

The heuristic utilizes Peer-2-Peer (P2P) [11] ideas centered on query routing, for identifying suitable neighboring domains which may contain the required resources. P2P and Grid systems share many properties for instance, it involve resource sharing across different administrative domains to support particular application behaviors. In P2P systems, a variety of mechanisms are available to achieve this, ranging from network flooding, more constrained gossiping protocols, to the use of structured overlays based on distributed hash tables. The demerit of the heuristics for inter domain meta-scheduling is the system will be overloaded due to bulk submission of large jobs. The nodes forward job queries o all neighbors using routing indices approach, but the peer to peer approach uses physics topology, which is no providing most efficient query forwarding which influence the scalability.

Mobile Agents

In grid resources from the Resource Brokers, the mobile agents are migrated to start sensors to measure the network metrics. Through the set of internal cost functions the raw data provided by the monitoring tools is used to produce a high level view of the Grid. The mobile agents are reduce the network load because they use less bandwidth by moving logic near data, and their actions are dependent on the state of the host environment. The mobile agents are capable of working without a dynamic connection between nodes; hence, it is not affected by network failures. So the mobile agents are extensively used in monitoring applications for information retrieval and resource discovery, and also for monitoring the network performance.

III. CONCLUSIONS

The survey helps in understanding the monitoring methodologies that are available for resource metrics and network metrics in grid systems according to their method categorization. With the numerous amounts of resource discovery techniques presented in this paper a clear idea is provided for how the resources are monitor in different architecture. The selection of resource Monitoring approach depends on the GRID environment.

Table-1: Comparison of Referred Papers

References	Methods Used	Method Description	Merits	Demerits
[1-3]	Relational Grid Monitoring Architecture (R-GMA).	To provide a flexible information & monitoring service for use by other middleware components And applications.	It allows queries to explore complex relationships in data.	The performance of R-GMA stability was hardening the code.
[4]	Net logger method.	Ability to change the logging level on the system.	It is very efficient and easy to use.	Ability to monitor & manage distributed computing component is critical for enabling high performance distributed computing.
[5-7]	Java Agents for Monitoring and Management (JAMM).	It sensors to filter the incoming events according to the consumer queries.	It used to monitor all system components and verify all required hardware and software was running properly.	It does not support the replication process.
[8]	Simple Network Management Protocol (SNMP).	To provide a "simple" method of centralizing the management of TCP/IP-based networks.	It is easy to use, easy to implement and easy to understand.	The problem of bandwidth limitation results from a huge number of requests (SNMP manager) and responses (SNMP agent).
[9-10]	Data Intensive and Network Aware (DIANA).	Able to consider network characteristics for scheduling decision making in the Grid.	It is a easy to use, easy to set up and not very difficult to understand.	Once jobs gets a CPU, then abort that job and Move that job to Grid site is difficult.
[11]	Inter domain Inter-domain Meta-scheduling.	To determine where a request for resources should be forwarded, akin to the approach adopted in P2P systems.	Nodes will forward job queries to all neighbors using Peer-to-Peer Process.	Observe performance problem such as unexpectedly low throughput or high latency.

Sources: Authors Compilation.

IV. REFERENCES

[1] Maozhen, Li, and Baker, Mark, (2005). "The Grid Core Technologies", John Wiley & Sons.

[2] Agustín, Caminero; Omer, Rana; Blanca, Caminero, and Carmen, Carrión, (2011). "Network-Aware Heuristics for Inter-Domain Meta-Scheduling in Grids", Journal of Computer and System Sciences, Volume: 77, Number: 2, pp. 262-281, March.

[3] Cooke, A., Gray, A., and Nutt, W., "Relational Grid Monitoring Architecture: Mediating Information about The Grid", Journal of Grid Computing. Volume: 2, Issue: 4, pp. 323-339.

[4] Gunter, Dan; Brian L. Tierney; Craig, E. Tull, and Virmani, Vibha, (2003). "On-Demand Grid Application Tuning and Debugging with the Net Logger Activation Service", Proceedings of the 4th International Workshop on Grid Computing, November 17, pp.76.

[5] Lowekamp, Bruce; Tierney, Brian; Cottrell, Less; Richard Hughes-Jones; Kielmann, Thilo, and Martin Swany, (2004). "A hierarchy of Network Performance Characteristics for Grid Applications and Services, GFD-R-P.023, Proposed Recommendation.

[6] Foster, I., and Kesselman, C., "Globus: a Met computing Infrastructure Toolkit", International Journal of Supercomputing Applications and High Performance Computing, Volume: 1, Issue: 2, pp. 115-128.

[7] Valliyammai, C., and Thamarai Selvi S., (2008). "Relational Network Monitoring System for Grid Performance Optimization", presented in International Conference ADCOM' 2008, pp. 170-173.

[8] Kwang, Sik Shin, Jin Ha Jung , Jin Young Cheon, Sang Bang Choi, "Real-time network monitoring scheme based on SNMP for dynamic information",

[9] Valliyammai, C.; Selvi, S.; Thamarai, Aravindh, S.; Jacob, Jervin, and Karthick, K., "Efficient Management of Network Resources Using Autonomous Agents in Grid", IJCSNS International Journal of Computer Science and Network Security. Volume: 8, Issue 4, pp. 178-184.

[10] Richard, M.; Anjum, A.; Stockinger, H.; Ali, A.; Willers, I., and Thomas, M., "Data Intensive and Network Aware (DIANA) Grid Scheduling", Journal Grid Computing. Volume: 5, Issue: 1, pp. 43-64.

[11] Rich, Wolski; Neil, T. Spring, and Hayes, Jim, (1999). "The Network Weather Service: a Distributed Resource Performance Forecasting Service for Met computing", Future Generation Computer Systems, October, Volume: 15, Number: 5 - 6, pp.757-768.

*****

DATA-PROVENANCE VERIFICATION FOR SECURE HOSTS

Sunitha R.[4] Senthilkumar[5]

ABSTRACT

Network or host-based signature scanning approaches alone were proven inadequate against new and emerging malware. We view malicious bots or malware in general as entities stealthily residing on a human user's computer and interacting with the user's computing resources. In this existing work we need to improve the trustworthiness of a host and its system data. Specifically, we provide a new mechanism that ensures the correct origin or provenance of critical system information and prevents adversaries from utilizing host resources. We define data-provenance integrity as the security property stating that the source where a piece of data is generated cannot be spoofed or tampered with. We describe a cryptographic provenance verification approach for ensuring system properties and system-data integrity at kernel-level. Its two concrete applications are demonstrated in the keystroke integrity verification and malicious traffic detection. Specifically, we first design and implement an efficient cryptographic protocol. The protocol prevents the forgery of fake key events by malware under reasonable assumptions. Then, we demonstrate our provenance verification approach by realizing a lightweight framework for restricting outbound malware traffic. We propose a malware detection approach based on the characteristic behaviors of human users. We explore the human-malware differences and classifies them to aid the detection of infected hosts by using support vector Machine (SVM) models. The server can collects the keystroke of particular client and classifies using SVM then verifies the data provenance for secure host. The existing Cryptographic provenance verification fails to detect the malware or trustworthiness of client in their own machine means it can be easily finds the trustworthy by the server using the classification models. There are two main challenges in this proposed work: one is how to select characteristic behavior features for classification, and the other is how to prevent malware forgeries.

KEYWORDS

SVM Model, Protocaol, Data, Provience, Hosts etc.

I. INTRODUCTION

Today, a security-conscious user who wants to verify that her input is not observed by malicious code during a sensitive online financial transaction faces an impasse. Key-loggers can capture a user's typed input and screen scrapers can process the content displayed to the user to obtain sensitive information such as credit card numbers. These malware exploit the vulnerabilities that are endemic to the huge computing base that is trusted to secure our private information.

Today's popular operating systems employ monolithic kernels, meaning that a vulnerability in any part of OS renders users' sensitive data insecure regardless of what application they may be running. On top of this untrustworthy OS sits a complex and monolithic web browser, which faces protection and assurance challenges similar to those of OS. It is not surprising that trusting this software stack for the protection of private data in web transactions often leads to data compromise.

Keystroke-Dynamics

Keystroke-dynamics based authentication is a cheap biometric mechanism that has been proven accurate in distinguishing individuals in much research. Most of the attack models considered in keystroke-dynamics literature assumes the attackers are humans, e.g., a colleague of Alice trying to log in as Alice. However, there has not been much study on the robustness of this technique against synthetic and automatic attacks and forgeries. For example, an attacker may write a program that performs statistic manipulation and synthesis to produce keystroke sequences in order to spoof others. These types of forgery attacks pose a serious threat. It is unclear from the current literature how robust keystroke dynamics is against forgery attacks.

Authentication versus Identification

Keystroke dynamics is part of a larger class of biometrics known as behavioral biometrics; their patterns are statistical in nature. It is a commonly held belief that behavioral biometrics are not reliable as physical biometrics used for authentication such as fingerprints or retinal scans or DNA. The reality here is that behavioral biometrics use a *confidence* measurement instead of the traditional *pass/fail* measurements. As such, the traditional benchmarks of False Acceptance Rate (FAR) and False Rejection Rates (FRR) no longer have linear relationships. The benefit to keystroke dynamics (as well as other behavioral biometrics) is that FRR/FAR can be adjusted by changing the acceptance threshold **at the individual level**.

This allows for explicitly defined individual risk mitigation–something physical biometric technologies could never achieve. Another benefit of keystroke dynamics: they can be captured continuously not just at the start-up time and may be adequately accurate to trigger an alarm to another system or person to come double-check the situation. In some cases, a person at gun-point might be forced to get start-up access by entering a password or having a particular fingerprint, but then that person could be replaced by someone else at the keyboard who was taking over for some bad purpose. In other less dramatic cases, an employee might violate business rules by sharing his password with his secretary, or by logging onto a system but then leaving the computer logged-in while someone else he knows about or doesn't know about uses the system.

Keystroke dynamics is one way to detect such problems sufficiently reliably to be worth investigating, because even a 20% true-positive rate would send the word out that this type of behavior is being watched and caught. Researchers are still a long way from being able to read a key logger session from a public computer in a library or cafe somewhere and identify the person from the keystroke dynamics, but we may be in a position to confidently **rule out** certain people from being the author, who we are confident is "a left-handed person with small hands who doesn't write in English as their primary language."

[4]Research Scholar, Karpagam Institute of Technology, Tamil Nadu, India, suni_pgt@yahoo.com

[5] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, rushtosen19@gmail.com

Temporal Variation

One of the major problems that keystroke dynamics runs into is that a person's typing varies substantially during a day and between different days. People may get tired, or angry, or have a beer, or switch computers, or move their keyboard tray to a new location, or use a virtual keyboard, or be pasting in information from another source (cut-and-paste), or from a voice-to-text converter. Even while typing, a person, for example, may be on the phone or pausing to talk. And some mornings, perhaps after a long night with little sleep and a lot of drinking, a person's typing may bear little resemblance to the way he types when he is well-rested. Extra doses of medication or missed doses could change his rhythm. There are hundreds of confounding circumstances. Because of these variations, any system will make false-positive and false-negative errors. Some of the successful commercial products have strategies to handle these issues and have proven effective in large-scale use (thousands of users) in real-world settings and applications.

II. APPLICATIONS OF DOMAIN

Our research domain focuses on a host-based approach for ensuring system-level data integrity and demonstrates its application for malware detection. In comparison, network trace analysis typically characterizes malware communication behaviors for detection. Such solutions usually involve pattern-recognition and machine learning techniques, and have demonstrated effectiveness against today's malware. Traces of botnets' command-and-control (C&C) messages i.e., how bots communicate with their botmasters are captured and their signatures and access patterns analyzed. For example, a host may be infected if it periodically contacts a server via the IRC (Internet relay chat) protocol and sends a large number of emails afterward.

Network traffic analysis can be realized by local Internet service providers to monitor and screen a large number of hosts as part of a network intrusion-detection system. To enforce the integrity of the detection systems, a virtual machine monitor (VMM) is usually required in particular for root kit detection. In this work, we utilize the existing trusted computing infrastructure. TPM is available on most commodity computers. The advantage of using TPM in comparison to VMM is the ease of accessing a host's kernel, and the ability to construct application-level fine-grained detection solutions, as described in the future work section.

How to address the runtime limitation of TPM is still an active research topic. For example, Flicker is a recently-proposed trusted computing base that allows sensitive applications to run in isolation in an untrusted operating system. In comparison, our trusted computing architecture supports functions beyond application integrity including enabling remote collection and verification of dynamic system data such as user-input events.

III. PROBLEM STATEMENT

Network- or host-based signature scanning approaches alone were proven inadequate against new and emerging malware. We view malicious bots or malware in general as entities stealthily residing on a human user's computer and interacting with the user's computing resources. For example, the malware may issue network calls to send outbound traffic for denial-of-service attacks, spam, or botnet command-and-control. However, conventional operating systems typically allow flexible execution pathways and data flow patterns, and are not specifically designed to distinguish legitimate user-initiated networking or file-system activities from malware triggered ones.

Our present work provides a hardware-based integrity service to address that problem. In comparison to which is designed specifically for browser input verification, our work provides a more general system-level solution for keystroke integrity that is application-oblivious.

III. OBJECTIVES OF PROJECT

Our objective is to improve the trustworthiness of the OS-level data flow; specifically, we provide mechanisms that ensure the correct origin or provenance of critical system data, which prevents adversaries from utilizing host resources (e.g., networking API).

We define a new security property - data-provenance integrity. It states that the source from which a piece of data is generated can be verified. We give the concrete illustration of how data-provenance integrity can be realized for system-level data namely, keystroke events and outbound network packets in a host- based setting.

IV. SCOPE OF PROJECT

Our scope of the project is to improve the assurance of system data and properties of a host, which has applications in preventing and identifying malware activities.

Our host-based system security solutions against malware complement network-traffic-based analysis. We demonstrated cryptographic provenance verification (CPV) application in identifying stealthy malware activities of a host, in particular how to distinguish malicious/unauthorized data flow from legitimate one on a computer that may be compromised.

We develop a TPM system that sign and verify keystroke events that are from external keyboard devices in client-server architecture, i.e., verifying the provenance of keystrokes. Our application of this system for distinguishing user inputs from malware inputs, which is useful in many scenarios including keystroke-dynamics based authentication. Our method has general application beyond the specific keystroke and network traffic problems studied in literature.

V. MOTIVATION

Authentication is the act of confirming the truth of an attribute of a datum or entity. This involve confirming the identity of a person, tracing the origins of an artifact, ensuring that a product is what its packaging and labeling claims to be, or assuring that a computer program is a trusted one. An identity can be authenticated in three ways: by something the user knows (such as a password or personal identification number), something the user has (a security token or smart card) or something the user is (a physical characteristic, such as a fingerprint, called a biometric).

Authentication is the process of validating a user, in order to set privileges based on a policy. This process is often conducted before a session is created, using a username and password combination.

To strengthen security, implementations of biometrics can help the process of user validation. Authentication systems allow entities to access to controlled resources. Traditionally, individuals are used to authenticate themselves on computers by using a classical couple of login and password. Strong authentication has for objective to use multiple authenticators for security purposes. The biometric features are conveniently divided into two main categories.

The physiological features include face, eye, fingerprints, palm topology, hand geometry, wrist veins and thermal images. The behavioral features include voiceprints, handwritten signatures and keystroke dynamics. Unlike other biometric methods, keystroke analysis can be done without the aid of special tools, just the keyboard of the computer where the biometric analysis has to be performed.

Biometric features are interesting for computer security because, on the one hand, they are sufficiently unique to be used to recognize legal users of systems and to reject impostors and, on the other hand, they cannot be forgotten, lost, overheard, stolen or extorted. Keystroke dynamics is considered as a strong behavioral biometric based authentication system.

Keystroke dynamics is a process of analyzing the way a user types at a terminal by monitoring the keyboard in order to identify the users based on habitual typing rhythm patterns. Moreover, unlike other biometric systems, which may be expensive to implement, keystroke dynamics is almost free as the only hardware required is the keyboard.

There are two approaches in keystroke authentication: a) Static, and b) Dynamic. Static approach authenticates the user at logon time and Dynamic methods authenticates after logon. The major difference between static and continuous keystroke dynamics is in the authentication phase. In static keystroke dynamics the complete typing of the fixed text in the input and template is compared, be this cannot be applied to continuous keystroke dynamics. In continuous keystroke dynamics, confidence levels are used. At each point in time, we must determine how confident we are that the user has not changed, based upon previous typing behavior.

The following factors are some keystroke dynamic features: a) the time that the key is held in pressing mode, b) key pressing intervals, and c) total speed of typing etc. The keystroke dynamics can be analyzed by various methods like fuzzy logic techniques, statistical methods and neural networks. Also the features can be extracted from user typing styles such as the temperature and pressure of user fingers when pressing the keys and the timing pattern of keystrokes.

The quality of keystroke dynamics patterns can be defined in terms of two factors: a) uniqueness and b) consistency. A combination of a high consistency and a high uniqueness will lead to a better discriminability or the ability to make better classification of user's patterns and impostor's patterns.

Our goal in this work is to improve the trustworthiness of a host and its system data. Specifically, we provide a new mechanism that ensures the correct origin or provenance of critical system information and prevents adversaries from utilizing host resources. We define data-provenance integrity as the security property stating that the source where a piece of data is generated cannot be spoofed or tampered with. We describe a cryptographic provenance verification approach for ensuring system properties and system-data integrity at kernel-level.

VI. CONCLUSIONS

We described a general approach for improving the assurance of system data and properties of a host, which has applications in preventing and identifying malware activities. Our host-based system security solutions against malware complement network-traffic-based analysis.

We demonstrated CPV's application in identifying stealthy malware activities of a host, in particular how to distinguish malicious/unauthorized data flow from legitimate one on a computer that may be compromised.

We made the following technical contributions in this paper:

1) We proposed the model and operations of cryptographic provenance verification in a host-based security setting,
2) We described an efficient keystroke integrity verification protocol in a client-server architecture that prevents malicious bots from forging keystroke events.
3) We explore the human-malware differences and classifies them to aid the detection of infected hosts by using support vector Machine (SVM) models.

VI. REFERENCES

[1] Baliga, Arati; Kamat, Pandurang, and Iftode, Liviu, "Lurking in the Shadows: Identifying Systemic Threats to Kernel Data".

[2] Sekar, R., "An Efficient Black-Box Technique for Defeating Web Application Attacks".

[3] Cui, Weidong; Randy H. Katz; and Wai-tian Tan, "Design and Implementation of an Extrusion-based Break-In Detector for Personal Computers".

[4] Stefan, Deian, and Danfeng, "Keystroke-Dynamics Authentication against Synthetic Forgeries".

[5] Senathipathi, K., and Krishnan, Batri, "Keystroke Dynamics Based Human Authentication System using Genetic Algorithm".

[6] Gang, Xu; Cristian, Borcea, and Liviu, Iftode, "Satem: Trusted Service Code Execution across Transactions".

[7] Xu, Kui, and Qiang, Ma, "Detecting Infection Onset with Behavior-Based Policies".

[8] Abhinav, Srivastava, and Jonathon, Giffin, "Tamper-Resistant, Application-Aware Blocking of Malicious Network Connections".

[9] Jonathan M., McCune, and Michael K. Reiter, "Safe Passage for Passwords and Other Sensitive Data".

*****

PERSUASIVE CUED CLICK-POINTS: DESIGN, IMPLEMENTATION & EVALUATION OF KNOWLEDGE BASED AUTHENTICATION MECHANISM

Ameefa P. K.[6] N. Mookhambika[7]

ABSTRACT

Click Cued Points is a click-based graphical password scheme, a cued-recall graphical password technique. Users Click on one point per image for a sequence of images. The next image is based on the previous click-point. Performance was very good in terms of speed, accuracy, and number of errors. Users preferred CCP to Pass Point, saying that selecting and remembering only one point per image was easier, and that seeing each image triggered their memory of where the corresponding point was located. CCP also provides greater security than Pass Points because the number of images increases the workload for attackers. Various graphical password schemes have been proposed as alternatives to text-based passwords. Research and experience have shown that text-based passwords are fraught with both usability and security problems that make them less than desirable solutions. Psychology studies have revealed that the human brain is better at recognizing and recalling images than text. Graphical passwords are intended to capitalize on this human characteristic in hopes that by reducing the memory burden on users, coupled with a larger full password space offered by images, more secure passwords can be produced and users will not resort to unsafe practices in order to cope. We propose a new click-based graphical password scheme called Cued Click Points (CCP). It can be viewed as a combination of Pass Points, Pass faces, and Story. A password consists of one click-point per image for a sequence of images. The next image displayed is based on the previous click-point so users receive immediate implicit feedback as to whether they are on the correct path when logging in. CCP offers both improved usability and security. Users could quickly create and re-enter their passwords. Another feature of CCP is the immediate implicit feedback telling the correct user whether their latest click-point was correctly entered.

KEYWORDS

Click Cued Points, Passwords, Security etc.

I. INTRODUCTION

Normally, Passwords are used for:

- Authentication (Establishes that the user is who they say they are),
- Authorization (The process used to decide if the authenticated person is allowed to access specific information or functions), and
- Access Control (Restriction of access-includes authentication & authorization).

In Figure-1, a graphical password system with a supportive sound signature to increase the remembrance of the password is discussed. In proposed work a click-based graphical password scheme called Cued Click Points (CCP) is presented. In this system a password consists of sequence of some images in which user can select one click-point per image. In addition user is asked to select a sound signature corresponding to each click point this sound signature will be used to help the user in recalling the click point on an image. System showed very good Performance in terms of speed, accuracy, and ease of use. Users preferred CCP to Pass Points, saying that selecting and remembering only one point per image was easier and sound signature helps considerably in recalling the click points. Cued Click Points (CCP) is a proposed alternative to PassPoints. In CCP, users click one point on each of 5 images rather than on five points on one image. It offers cued-recall and introduces visual cues that instantly alert valid users if they have made a mistake when entering their latest click-point (at which point they can cancel their attempt and retry from the beginning). It also makes attacks based on hotspot analysis more challenging.

Authentication

The only significant user study on the security of graphical passwords for authentication was performed by Davis. Present authors.23 in that work; we studied the security of two schemes based on image recognition, denoted "Face" and "Story," which are described shortly. This study focused specifically on the impact of user selection of passwords in these schemes, and the security of the passwords that resulted. We recount some of the notable results from this study, and the methodologies used to reach them, as an illustration of some of the challenges that graphical passwords can face. In particular, this study demonstrated that graphical password schemes can be far weaker than textual passwords when users are permitted to choose their password.

System Tolerance

After creation of the login vector, system calculates the Euclidian distance between login vector and profile vectors stored. Euclidian distance between two vectors '**p'** and '**q'** is given by-Above distance is calculated for each image if this distance comes out less than a tolerance value D. The value of D is decided according to the application. In our system this value is selected by the user.

Security

Security is the degree of protection to safeguard a nation, union of nations, persons or person against danger, damage, loss, and crime. Security as a form of protection is structures and processes that provide or improve security as a condition. The Institute for Security and Open Methodologies (ISECOM) in the OSSTMM 3 defines security as "a form of protection where a separation is created between the assets and the threat". This includes but is not limited to the elimination of either the asset or the threat. Security as a national condition was defined in a United Nations study (1986) so that countries can develop and progress safely. Security has to be compared

[6] Student, Final Year, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, affa2006@gmail.com
[7] Assitant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

to related concepts: safety, continuity, reliability. The key difference between security and reliability is that security must take into account the actions of people attempting to cause destruction. Different scenarios also give rise to the context in which security is maintained:

- With respect to classified matter, the condition that prevents unauthorized persons from having access to official information that is safeguarded in the interests of national security.
- Measures taken by a military unit, an activity or installation to protect itself against all acts designed to, or which may, impair its effectiveness.

Security Concepts

Certain concepts recur throughout different fields of security:

- Assurance – It is the level of guarantee that a security system will behave as expected,
- Countermeasure – It is a way to stop a threat from triggering a risk event,
- Defense in depth - never rely on one single security measure alone,
- Exploit - a vulnerability that has been triggered by a threat - a risk of 1.0 (100%),
- Risk – is a possible event which could cause a loss,
- Threat - is a method of triggering a risk event that is dangerous,
- Vulnerability - a weakness in a target that can potentially be exploited by a security threat.

Security Management in Organizations

In the corporate world, various aspects of security were historically addressed separately - notably by distinct and often non communicating departments for IT security, physical security, and fraud prevention. Today there is a greater recognition of interconnected nature of security requirements, an approach variously known as holistic security, "all hazards" management, and other terms. Inciting factors in the convergence of security disciplines include the development of digital video surveillance technologies (see Professional video over IP) and the digitization and networking of physical control systems (see SCADA). Greater interdisciplinary cooperation is further evidenced by the February 2005 creation of the Alliance for Enterprise Security Risk Management, a joint venture including leading associations in security (ASIS), information security (ISSA, the Information Systems Security Association), and IT audit (ISACA, the Information Systems Audit and Control Association). In 2007 the International Organisation for Standardization (ISO) released ISO 28000 - Security Management Systems for the supply chain. Although the title supply chain is included, this Standard specifies the requirements for a security management system, including those aspects critical to security assurance for any organisation or enterprise wishing to management the security of the organisation and its activities. ISO 28000 is the foremost risk based security system and is suitable for managing both public and private regulatory security, customs and industry based security schemes and requirements.

Authentication

Authentication is the act of confirming the truth of an attribute of a datum or entity. This might involve confirming the identity of a person or software program, tracing the origins of an artifact, or ensuring that a product is what it's packaging and labeling claims to be.

Authentication Methods

In art, antiques, and anthropology, a common problem is verifying that a person has the said identity, or a given artifact was produced by a certain person or was produced in a certain place or period of history.

There are three types of techniques:

The **first type** of authentication is accepting proof of identity given by a credible person who has evidence on the said identity, or on the originator and the object under assessment as the originator's artifact respectively. The **second type** of authentication is comparing the attributes of the object itself to what is known about objects of that origin. For example, an art expert might look for similarities in the style of painting, check the location and form of a signature, or compare the object to an old photograph. An archaeologist might use carbon dating to verify the age of an artifact, do a chemical analysis of the materials used, or compare the style of construction or decoration to other artifacts of similar origin. The physics of sound and light, and comparison with a known physical environment, can be used to examine the authenticity of audio recordings, photographs, or videos. Attribute comparison may be vulnerable to forgery. In general, it relies on the facts that creating a forgery indistinguishable from a genuine artifact requires expert knowledge, that mistakes are easily made, and that the amount of effort required to do so is considerably greater than the amount of profit that can be gained from the forgery. In art and antiques, certificates are of great importance for authenticating an object of interest and value. Certificates can, however, also be forged, and the authentication of these poses a problem. For instance, the son of Han van Meegeren, the well-known art-forger, forged the work of his father and provided a certificate for its provenance as well; see the article Jacques van Meegeren. Criminal and civil penalties for fraud, forgery, and counterfeiting can reduce the incentive for falsification, depending on the risk of getting caught. The **third type** of authentication relies on documentation or other external affirmations. For example, the rules of evidence in criminal courts often require establishing the chain of custody of evidence presented. This can be accomplished through a written evidence log, or by testimony from the police detectives and forensics staff that handled it. Some antiques are accompanied by certificates attesting to their authenticity. External records have their own problems of forgery and perjury, and are also vulnerable to being separated from the artifact and lost.

Currency and other financial instruments commonly use the first type of authentication method. Bills, coins, and cheques incorporate hard-to-duplicate physical features, such as fine printing or engraving, distinctive feel, watermarks, and holographic imagery, which are easy for receivers to verify. Consumer goods such as pharmaceuticals, perfume, fashion clothing can use either type of authentication method to prevent counterfeit goods from taking advantage of a popular brand's reputation (damaging the brand owner's sales and reputation). A trademark is a legally protected marking or other identifying feature which aids consumers in the identification of genuine brand-name goods.

II. RELATED WORKS

Performance of Hase-Based Quantum Key Two Server user Password Authentication for Internet Services: The authentication system uses passwords to authenticate user login and stores the password in a central server. The central server is easily prone to attack and if they are being compromised by the intruder, it is possible for the intruder to obtain the password and gain access to the contents of the user. To overcome this problem, the multi-server systems were being proposed in which the user has to communicate in parallel with several or all of the servers for the purpose of authentication. Multi-server system requires a large communication bandwidth and needs for synchronization at the user.

Visualization of Textual Passwords: Passwords are now everywhere. The main form of passwords is based on characters you can type on your keyboard, normally called textual passwords. One major security problem with textual passwords is its vulnerability to dictionary attack, namely, brute-force attack based on a dictionary which is much smaller than the whole password space. In this project, you will develop an interactive program to visualize the security of a textual password w.r.t. one or more given dictionaries, and to help the user to select a more secure textual password while he/she is typing the password.

Attacks against Textual Passwords: Attackers generally compromise passwords in one of four ways:

- By gathering enough information about users to guess their password;
- By social engineering, e.g., tricking users into revealing their usernames and/or passwords;
- By capturing users' passwords, e.g., via shoulder surfing or spyware
- By cracking passwords using a software program, such as John the Ripper.

Human Selection of Mnemonic Phrase-based Passwords: Textual passwords are often the only mechanism used to authenticate users of a networked system. Unfortunately, many passwords are easily guessed or cracked. In an attempt to strengthen passwords, some systems instruct users to create mnemonic phrase-based passwords. A mnemonic password is one where a user chooses a memorable phrase and uses a character (often the first letter) to represent each word in the phrase. In this paper, we hypothesize that users will select mnemonic phrases that are commonly available on the Internet, and that it is possible to build a dictionary to crack mnemonic phrase-based passwords.

Picture Password: A Visual Login Technique for Mobile Devices Adequate user authentication is a persistent problem, particularly with handheld devices such as Personal Digital Assistants (PDAs), which tend to be highly personal and at the fringes of an organization's influence. Yet, these devices are being used increasingly in corporate settings where they pose a security risk, not only by containing sensitive information, but also by providing the means to access such information over wireless network interfaces. User authentication is the first line of defense for a lost or stolen PDA. However, motivating users to enable simple PIN or password mechanisms and periodically update their authentication information is a constant struggle.

III. PROPOSED SYSTEM

- Click cued points is a click-based graphical password scheme, a cued-recall graphical password technique.
- Various graphical password schemes have been proposed as alternatives to text-based passwords.
- It can be used as password for folder lock, web-driven applications, desktop lock etc.
- In proposed work a click-based graphical password scheme called Cued Click Points (CCP) is presented. In this system a password consists of sequence of some images in which user can select one click-point per image.

Cued Click Points (CCP) is a proposed alternative to PassPoints. In CCP, users click point on images rather than on five points on one image. It offers cued-recall and introduces visual cues that instantly alert valid users if they have made a mistake when entering their latest click-point (at which point they can cancel their attempt and retry from the beginning). It also makes attacks based on hotspot analysis more challenging. Each click results in showing a next-image, in effect leading users down a "path" as they click on their sequence of points. A wrong click leads down an incorrect path, with an explicit indication of authentication failure only after the final click. Users can choose their images only to the extent that their click-point dictates the next image. If they dislike the resulting images, they could create a new password involving different click-points to get different images.

- System showed very good Performance in terms of speed, accuracy, and ease of use. Users preferred CCP to Pass Points, saying that selecting and remembering only one point per image was easier and sound signature helps considerably in recalling the click points.
- In the proposed work we have integrated sound signature to help in recalling the password. No system has been devolved so far which uses sound signature in graphical password authentication.

IV. ARCHIETECTURE

Problem Statement

The password techniques recently used in market are very insecure. Theses password used are easily hacked by the intruders from the network. The textual password used is easily guessed. To overcome these problems the market was provided with techniques like OTP (One Time Password). But the OTP password is provide by token devices. These token devices are very expensive.

Project Scope

Cued Click Points (CCP) is a proposed alternative to PassPoints. In CCP, users click point on images rather than on five points on one image. It offers cued-recall and introduces visual cues that instantly alert valid users if they have made a mistake when entering their latest click-point (at which point they can cancel their attempt and retry from the beginning). It also makes attacks based on hotspot analysis more challenging. Each click results in showing a next-image, in effect leading users down a "path" as they click on their sequence of points.

A wrong click leads down an incorrect path, with an explicit indication of authentication failure only after the final click.

Users can choose their images only to the extent that their click-point dictates the next image. If they dislike the resulting images, they could create a new password involving different click-points to get different images.

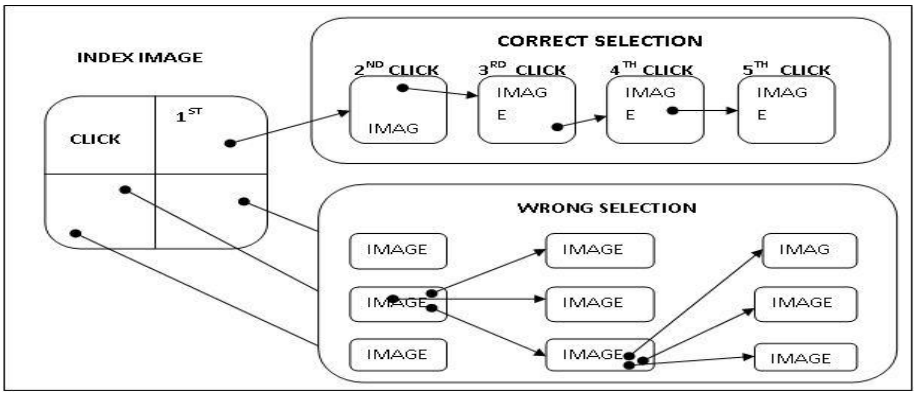
- ***Registration***

  In PassPoints, passwords consist of a sequence of five click-points on a given image. Users may select any pixels in image as click-points for their password. To log in, they repeat the sequence of clicks in correct order, within a system-defined tolerance square of the original click-points.

- ***Viewport Creation***

  By adding a persuasive feature to CCP encourages users to select less predictable passwords, and makes it more difficult to select passwords where all five click-points are hotspots. Specifically, when users create a password, the images are slightly shaded except for a viewport. The viewport is positioned randomly, rather than specifically to avoid known hotspots, since such information might allow attackers to improve guesses and could lead to the formation of new hotspots. The viewport's size is intended to offer a variety of distinct points but still cover only an acceptably small fraction of all possible points.

Figure-1



Sources: Authors Compilation.

CCP Management

Users must select a click-point within this highlighted viewport and cannot click outside of the viewport, unless they press the shuffle button to randomly reposition the viewport. While users may shuffle as often as desired, this significantly slows password creation. The viewport and shuffle button appears only during password creation. During later password entry, the images are displayed normally, without shading or the viewport, and users may click anywhere on the images.

Authentication

The password entry becomes a true cued-recall scenario, wherein each image triggers the memory of a corresponding click-point. Remembering the order of the click-points is no longer a requirement on users, as the system presents the images one at a time. CCP also provides implicit feedback claimed to be useful only to legitimate users. When logging on, seeing an image they do not recognize alerts users that their previous click-point was incorrect and users may restart password entry. Explicit indication of authentication failure is only provided after the final click-point, to protect against incremental guessing attacks.

IV. CONCLUSIONS

In this project user should remember the click points on image. Also user need to upload the images by own**.** This study focused specifically on the impact of user selection of passwords in these schemes, and the security of the passwords that resulted. We recount some of the notable results from this study, and the methodologies used to reach them, as an illustration of some of the challenges that graphical passwords can face. In particular, this study demonstrated that graphical password schemes can be far weaker than textual passwords when users are permitted to choose their password. After creation of the login vector, system calculates the Euclidian distance between login vector and profile vectors stored. Euclidian distance between two vectors is denated as **p** and **q**. Above distance is calculated for each image if this distance comes out less than a tolerance value D. The value of D is decided according to the application. In our system this value is selected by the user.

V. REFERENCES

[1] Birget, J. C.; D. Hong, and N. Memon, (2006). "Graphical Passwords Based on Robust Discretization", IEEE Trans. Info. Forensics and Security, September, 1(3).

[2] Blonder, G. E., (1996). "Graphical Passwords", United States Patent: 5, 559, 961.

[3] Chiasson, S.; Biddle; R., and P. C. van Oorschot, (2007). "Second Look at Usability of Click - Based Graphical Passwords", ACM SOUPS

[4] Cranor, L. F., and S. Garfinkel, (2005). "Security and Usability", O' Reilly Media.

[5] Davis, D.; F. Monrose, and M. K. Reiter, (2004). "On User Choice in Graphical Password Schemes", 13th USENIX Security Symposium.

*****

PRIVACY PRESERVING DECISION TREE LEARNING USING UNREALIZED DATA SETS

Beula Amalorpavam A.[8] N. Mookhambika[9]

ABSTRACT

Privacy preservation is important for machine learning and data mining, but measures designed to protect private information often result in a trade-off: reduced utility of training samples. This paper introduces a privacy preserving approach that can be applied to decision tree learning, without concomitant loss of accuracy. It describes an approach to the preservation of the privacy of collected data samples in cases where information from the sample database has been partially lost. This approach converts the original sample data sets into a group of unreal data sets, from which the original samples cannot be reconstructed without the entire group of unreal data sets. Meanwhile, an accurate decision tree can be built directly from those unreal data sets. This novel approach can be applied directly to the data storage as soon as the first sample is collected. The approach is compatible with other privacy preserving approaches, such as cryptography, for extra protection.

KEYWORDS

Data, Privacy, Preservation, Decision Tree etc.

I. INTRODUCTION

DATA mining is widely used by researchers for science and business purposes. Data collected (referred to as "sample data sets" or "samples") from individuals (referred to "information providers") are important for decision making or pattern recognition. Therefore, privacy-preserving processes have been developed to sanitize private information from the samples while keeping their utility.

A large body of research has been devoted to the protection of sensitive information when samples are given to third parties for processing or computing [1], [2], [3], [4], [5]. It is in the interest of research to disseminate samples to a wide audience of researchers, without making strong assumptions about their trustworthiness. Even if information collectors ensure that data are released only to third parties with non malicious intent (or if a privacy preserving approach can be applied before the data are released,), there is always the possibility that the information collectors may inadvertently disclose samples to malicious parties or that the samples are actively stolen from the collectors. Samples may be leaked or stolen anytime during the storing process [6], [7] or while residing in storage [8], [9]. This paper focuses on preventing such attacks on third parties for the whole lifetime of the samples.

Contemporary research in privacy preserving data mining mainly falls into one of two categories: 1) perturbation and randomization-based approaches, and 2) secure multiparty computation (SMC)-based approaches [10].SMC approaches employ cryptographic tools for collaborative data mining computation by multiple parties. Samples are distributed among different parties and they take part in the information computation and communication process. SMC research focuses on protocol development for protecting privacy among the involved parties or computation efficiency; however, centralized processing of samples and storage privacy is out of the scope of SMC. We introduce a new perturbation and randomization based approach that protects centralized sample data sets utilized for decision tree data mining. Privacy preservation is applied to sanitize the samples prior to their release to third parties in order to mitigate the threat of their inadvertent disclosure or theft. In contrast to other sanitization methods, our approach does not affect the accuracy of data mining results. The decision tree can be built directly from the sanitized data sets, such that the originals do not need to be reconstructed. Moreover, this approach can be applied at any time during the data collection process so that privacy protection can be in effect even while samples are still being collected. The following assumptions are made for the scope of this paper: first, as is the norm in data collection processes, a sufficiently large number of sample data sets have been collected to achieve significant data mining results covering the whole research target. Second, the number of data sets leaked to potential attackers constitutes a small portion of the entire sample database. Third, identity attributes (e.g., social insurance number) are not considered for the data mining process because such attributes are not meaningful for decision making. Fourth, all data collected are discretized; continuous values can be represented via ranged value attributes for decision tree data mining.

In Privacy Preserving Data Mining: Models and Algorithms [14], Aggarwal and Yu classify privacy preserving data mining techniques, including data modification and cryptographic, statistical, query auditing and perturbation-based strategies. Statistical, query auditing and most cryptographic techniques are subjects beyond the focus of this paper. In this section, we explore the privacy preservation techniques for storage privacy attacks.

Data modification techniques maintain privacy by modifying attribute values of the sample data sets. Essentially, data sets are modified by eliminating or unifying uncommon elements among all data sets. These similar data sets act as masks for the others within the group because they cannot be distinguished from the others; every data set is loosely linked with a certain number of information providers. k-anonymity [15] is a data modification approach that aims to protect private information of the samples by generalizing attributes. k-anonymity trades privacy for utility. Further, this approach can be applied only after the entire data collection process has been completed.

Perturbation-based approaches attempt to achieve privacy protection by distorting information from the original data sets. The perturbed data sets still retain features of the originals so that they can be used to perform data mining directly or indirectly via data reconstruction. Random substitutions [16] is a perturbation approach that randomly substitutes the values of selected attributes to achieve privacy protection for those attributes, and then applies data reconstruction when these data sets are needed for data mining. Even though privacy of the

[8] Student, Final Year, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, affa2006@gmail.com
[9] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

selected attributes can be protected, the utility is not recoverable because the reconstructed data sets are random estimations of the originals.

Most cryptographic techniques are derived for secure multiparty computation, but only some of them are applicable to our scenario. To preserve private information samples are encrypted by a function, f, (or a set of functions) with a key, k, (or a set of keys); meanwhile, original information can be reconstructed by applying a decryption function, f_1, (or a set of functions) with the key, k, which raises the security issues of the decryption function(s) and the key(s). Building meaningful decision trees needs encrypted data to either be decrypted or interpreted in its encrypted form. The (anti)monotone framework [17] is designed to preserve both the privacy and the utility of the sample datasets used for decision tree data mining. This method applies a series of encrypting functions to sanitize the samples and decrypts them correspondingly for building the decision tree. However, this approach raises the security concerns about the encrypting and decrypting functions. In addition to protecting the input data of the data mining process, this approach also protects the output data, i.e., the generated decision tree. Still, this output data can normally be considered sanitized because it constitutes an aggregated result and does not belong to any individual information provider. In addition, this approach does not work well for discrete-valued attributes.

II. DATA SET COMPLEMENTATION APPROACH

In the following sections, we will work with sets that can contain multiple instances of the same element, i.e., with multisets (bags) rather than with sets as defined in the classical set theory. We begin this section by defining fundamental concepts (Section 2.1). We then introduce our data unrealization algorithm.

Most existing data mining approaches perform data mining tasks on a single data table. However, increasingly, data repositories such as financial data and medical records, amongst others, are stored in relational databases. The inability of applying traditional data mining techniques directly on such relational database thus poses a serious challenge. To address this issue, a number of researchers convert a relational database into one or more flat files and then apply traditional data mining algorithms. The above-mentioned process of transforming a relational database into one or more flat files usually involves aggregation. Aggregation functions such as maximum, minimum, average, standard deviation; count and sum are commonly used in such a flattening process.

Our research aims to address the following question: Is there a link between aggregation and possible privacy violations during relational database mining? In this research we investigate how, and if, applying aggregation functions will affect the privacy of a relational database, during supervised learning, or classification, where the target concept is known. To this end, we introduce the PBIRD (Privacy Breach Investigation in Relational Databases) methodology. The PBIRD methodology combines multi-view learning with feature selection, to discover the potentially *dangerous* sets of features as hidden within a database. Our approach creates a number of views, which consist of subsets of the data, with and without aggregation. Then, by identifying and investigating the set of selected features in each view, potential privacy breaches are detected. In this way, our PBIRD algorithm is able to discover those features that are correlated with the classification target that may also lead to revealing of sensitive information in the database.

Our experimental results show that aggregation functions do, indeed, change the correlation between attributes and the classification target. We show that with aggregation, we obtain a set of features which can be accurately linked to the classification target and used to predict (with high accuracy) the confidential information. On the other hand, the results show that, without aggregation we obtain another different set of potentially harmful features. By identifying the complete set of potentially dangerous attributes, the PBIRD ii methodology provides a solution where the database designers/owners can be warned, to subsequently perform necessary adjustments to protect the privacy of the relational database.

In our research, we also perform a comparative study to investigate the impact of aggregation on the classification accuracy and on the time required to build the models. Our results suggest that in the case where a database consists only of categorical data, aggregation should especially be used with caution. This is due to the fact that aggregation causes a decrease in overall accuracies of the resulting models. When the database contains mixed attributes, the results show that the accuracies without aggregation and with aggregation are comparable. However, even in such scenarios, schemas without aggregation tend to slightly outperform. With regard to the impact of aggregation on the model building time, the results show that, in general, the models constructed with aggregation require shorter building time. However, when the database is small and consists of nominal attributes with high cardinality, aggregation causes a slower model building time.

2.1 Universal Set and Data Set Complement

Definition 1

TU, the universal set of data table T, is a set containing a single instance of all possible data sets in data table T.

Definition 2

If TD is a subset of T and q is a positive integer, then q-multiple-of TD, denoted as qTD, is a set of data sets containing q instances of each data set in TD.

Definition 3

If k is a possible value of attribute a and l is a possible value of attribute b in T, denotes a subset of T that contains all data sets with attribute a equals k. Similarly, denotes a subset of T that contains all data sets with attribute a equals k and b equals l.

Definition 4

If TD is a subset of T, then the absolute complement of TD, denoted as TcD , is equal to TU _ TD, and a q-absolute-complement of TD, denoted as qTcD , is equal to qTU _ TD.

2.2 Unrealized Training Set

Traditionally, a training set, TS, is constructed by inserting sample data sets into a data table. However, a data set complementation approach, as presented in this paper, requires an extra data table, TP. TP is a perturbing set that generates

unreal data sets which are used for converting the sample data into an unrealized training set, T0. The algorithm for unrealizing the training set.

The problem of privacy-preserving data mining has become more important in recent years because of the increasing ability to store personal data about users, and the increasing sophistication of data mining algorithms to leverage this information. A number of techniques such as randomization and k-anonymity [1, 4, 16] have been suggested in recent years in order to perform privacy-preserving data mining. Furthermore, the problem has been discussed in multiple communities such as the database community, the statistical disclosure control community and cryptography community.

The key directions in the field of privacy-preserving data mining are as follows:

a) Privacy-Preserving Data Publishing: These techniques tend to study different transformation methods associated with privacy. These techniques include methods such as randomization [1], k-anonymity [16, 7], and l-diversity [11]. Another related issue is how the perturbed data can be used in conjunction with classical data mining methods such as association rule mining [15]. Other related problems include that of determining privacy-preserving methods to keep the underlying data useful (utility-based methods), or the problem of studying the different definitions of privacy, and how they compare in terms of effectiveness in different scenarios.

b) Changing Results of Data Mining Applications to Preserve Privacy: In many cases, the results of data mining applications such as association rule or classification rule mining can compromise the privacy of the data. This has spawned a field of privacy inwhich the results of data mining algorithms such as association rule mining are modified in order to preserve the privacy of the data. A classic example of such techniques are association rule hiding methods, in which some of the association rules are suppressed in order to preserve privacy.

c) Query Auditing: Such methods are akin to the previous case of modifying the results of data mining algorithms. Here, we are either modifying or restricting the results of queries. Methods for perturbing the output of queries are discussed in [8], whereas techniques for restricting queries are discussed in [9, 13].

d) Cryptographic Methods for Distributed Privacy: In many cases, the data may be distributed across multiple sites, and the owners of the data across these different sites may wish to compute a common function. In such cases, a variety of cryptographic protocols may be used in order to communicate among the different sites, so that secure function computation is possible without revealing sensitive information. A survey of such methods may be found in [14].

e) Theoretical Challenges in High Dimensionality: Real data sets are usually extremely high dimensional, and this makes the process of privacy preservation extremely difficult both from a computational and effectiveness point of view. In [12], it has been shown that optimal k-anonymization isNP-hard. Furthermore, the technique is not even effectivewith increasing dimensionality, since the data can typically be combined with either public or background information to reveal the identity of the underlying record owners.

The elements in the resulting data sets are unreal individually, but meaningful when they are used together to calculate the information required by a modified ID3 algorithm.

III. DECISION TREE GENERATION

The well-known ID3 algorithm [11] builds a decision tree by calling algorithm Choose-Attribute recursively. This algorithm selects a test attribute (with the smallest entropy) according to the information content of the training set TS and the algorithm Majority-Value retrieves the most frequent value of the decision attribute of TS.

3.1 Information Entropy Determination

From the algorithm Unrealize-Training-Set, it is obvious that the size of TS is the same as the size of T0. Furthermore, all data sets in (T0 þ TP) are based on the data sets in TU, excepting the ones in TS, i.e., TS is the q-absolutecomplement of (T0 þ TP) for some positive integer q. According to Theorem 2, the size of qTU can be computed from the sizes of T0 and TP, with qTU ¼ 2_jT0j þ jTP j. Therefore, entropies of the original data sets, TS, with any decision attribute and any test attribute, can be determined by the unreal training set, T0, and perturbing set, TP.

3.2 Modified Decision Tree Generation

Algorithm As entropies of the original data sets, TS, can be determined by the retrievable information—the contents of unrealized training set, T0, and perturbing set, TP—the decision tree of TS can be generated by following algorithm.

3.3 Data Set Reconstruction

Section 4.2 introduced a modified decision tree learning algorithm by using the unrealized training set, T', and the perturbing set, TP. Alternatively, we could have reconstructed the original sample data sets, TS, from T0 and TP (shown in Fig. 5), followed by an application of the conventional ID3 algorithm for generating the decision tree 4.4 Enhanced Protection with Dummy Values Dummy values can be added for any attribute such that the domain of the perturbed sample data sets will be expanded while the addition of dummy values will have no impact on TS. For example, we can expand the possible values of attribute Wind from {Strong, Weak} to {Dummy, Strong, Weak} where Dummy represents a dummy attribute value that plays no role in the data collection process In this way we can keep the same resulting decision tree (because the entropy of TS does not change) while arbitrarily expanding the size of TU. Meanwhile, all data sets in T0 and TP, including the ones with a dummy attribute value, are needed for determining the entropies of q½T0 þ TP _c during the decision tree generation process.

IV. THEORETICAL EVALUATION

This section provides a concise theoretical evaluation of our approach.

4.1 Privacy Issues

Private information could potentially be disclosed by the leaking of some sanitized data sets, TL (a subset of the entire collected data table, TD), to an unauthorized party if 1. The attacker is able to reconstruct an original sample, tS, from TL,

or 2. If tL (a data set in TL) matches tS (a data set in TS) by chance. In the scope of this paper, tS is non reconstructable because jTLj is much smaller than jT0 þ TP j. Hence, we are focusing on the privacy loss via matching. Without privacy preservation the collected data sets are the original samples. Samples with more even distribution (low variance) have less privacy loss, while data sets with high frequencies are at risk.

The data set complementation approach solves the privacy issues of those uneven samples. This approach converts the original samples into some unrealized data sets [T0 þ TP], such that the range of privacy loss.

Data Set complementation is in favor of those samples with high variance distribution, especially when some data sets have zero counts. However, it does not provide significant improvement for the even cases.

In Section 3.4, we showed that we can generate zero count data sets and "uneven" the data set distribution by adding dummy attribute values. If we expand the size of the possible sample domain then the range of privacy loss will be decreased.

4.2 Privacy-Preserving Data Mining Algorithms

In this section, we will discuss the key stream mining problems and will discuss the challenges associated with each problem.

The Randomization Method

The randomization technique uses data distortion methods in order to create private representations of the records [1, 4]. In most cases, the individual records cannot be recovered, but only aggregate distributions can be recovered. These aggregate distributions can be used for data mining purposes. Two kinds of perturbation are possible with the randomization method:

a) Additive Perturbation: In this case, randomized noise is added to the data records. The overall data distributions can be recovered from the randomized records. Data mining and management algorithms re designed to work with these data distributions.

b) Multiplicative Perturbation: In this case, the random projection or random rotation techniques are used in order to perturb the records.

V. EXPERIMENTS

This section shows the experimental results from application of the data set complementation approach to:

1. Normally distributed samples,
2. Evenly distributed samples,
3. Extremely unevenly distributed samples, and
4. Six sets of randomly picked samples.

Where (i) was generated without creating any dummy attribute values, and (ii) was generated by applying the dummy attribute technique to double the size of the sample domain. For the artificial samples, we will study the output accuracy (the similarity between the decision tree generated by the regular method and by the new approach), the storage complexity (the space required to store the unrealized samples based on the size of the original samples) and the privacy risk (the maximum, minimum, and average privacy loss if one unrealized data set is leaked). The unrealized samples, T0 and TP, are shown. For the random samples, we will study the storage complexity and the privacy risk of one data set loss and that of some portions (10, 20, 30, 40, and 50 percent of randomly picked data sets) of data set loss.

5.1 Output Accuracy

In all cases, the decision tree(s) generated from the unrealized samples (by algorithm Generate-Tree') is the same as the decision tree(s), TreeTs, generated from the original sample by the regular method.

5.2 Storage Complexity

From the experiment, the storage requirement for the data set complementation approach increases .while the required storage may be doubled if the dummy attribute values technique is applied to double the sample domain. The best case happens when the samples are evenly distributed, as the storage requirement is the same as for the originals. The worst case happens when the samples are distributed extremely unevenly. Based on the randomly picked tests, the storage requirement for our approach is less than five times (without dummy values) and eight times (with dummy values, doubling the sample domain) that of the original samples.

5.3 Privacy Risk

Without the dummy attribute values technique, the average privacy loss per leaked unrealized data set is small, except for the even distribution case (in which the unrealized samples are the same as the originals). By doubling the sample domain, the average privacy loss for a single leaked data set is zero, as the unrealized samples are not linked to any information provider. The randomly picked tests show that the data set complementation approach eliminates the privacy risk for most cases and always improves privacy security significantly when dummy values are used.

Privacy Attacks

It is useful to examine the different ways in which one can make adversarial attacks on privacy-transformed data. This helps in designing more effective privacy-transformation methods. Some examples of methods which can be used in order to attack the privacy of the underlying data include SVD-based methods, spectral filtering methods and background knowledge attacks.

Cryptographicmethods for Information Sharing & Privacy

In many cases, multiple parties may wish to share aggregate private data, without leaking any sensitive information at their end [14]. For example, different superstores with sensitive sales data may wish to coordinate among themselves in knowing aggregate trends without leaking the trends of their individual stores. This requires secure and cryptographic protocols for sharing the information across the different parties. The data may be distributed in two ways across different sites:

a) Horizontal Partitioning: In this case, the different sites may have different sets of records containing same attributes.

b) Vertical Partitioning: In this case, the different sites may have different attributes of the same sets of records.

VI. CONCLUSIONS

We introduced a new privacy preserving approach via data set complementation which confirms the utility of training data sets for decision tree learning. This approach converts the sample data sets, TS, into some unreal data sets (T0 þ TP) such that any original data set is not re constructable if an unauthorized party were to steal some portion of (T0 þ TP). Meanwhile, there remains only a low probability of random matching of any original data set to the stolen data sets, TL. The data set complementation approach ensures that the privacy loss where TU is the set of possible sample data sets. By creating dummy attribute values and expanding the size of sample domain to R times.

Privacy preservation via data set complementation fails if all training data sets are leaked because the data set reconstruction algorithm is generic. Therefore, further research is required to overcome this limitation. As it is very straightforward to apply a cryptographic privacy preserving approach, such as the (anti) monotone framework, along with data set complementation, this direction for future research could correct the above limitation.

This paper covers the application of this new privacy preserving approach with the ID3 decision tree learning algorithm and discrete-valued attributes only. Future research should develop the application scope for other algorithms, such as C4.5 and C5.0, and data mining methods with mixed discretely and continuously valued attributes. Furthermore, the data set complementation approach expands the sample storage size (in the worst case, the storage size equals ð2jTUj _ 1Þ_ jTSjÞ; therefore, further studies are needed into optimizing - 1) the storage size of the unrealized samples, and 2) the processing time when generating a decision tree from those samples.

VII. REFERENCES

[1] Ajmani, S.; Morris, R., and Liskov, B., (2001). "A Trusted Third - Party Computation Service", Technical Report MIT-LCS-TR-847, MIT.

[2] Wang, S. L., and Jafari, A., (2005). "Hiding Sensitive Predictive Association Rules", Proceeding IEEE International Conference Systems, Man and Cybernetics, pp. 164-169.

[3] Agrawal, R., and Srikant, R., (2000). "Privacy Preserving Data Mining," Proceeding ACM SIGMOD Conference Management of Data (SIGMOD '00), May, pp. 439-450.

[4] Ma, Q., and Deng, P., (2008). "Secure Multi-Party Protocols for Privacy Preserving Data Mining", Proceeding Third International Conference - Wireless Algorithms, Systems, and Applications (WASA '08), pp. 526-537.

[5] Gitanjali, J.; Indumathi, J.; Iyengar, N. C., and Sriman, N., (2010). "A Pristine Clean Cabalistic Foruity Strategize Based Approach for Incremental Data Stream Privacy Preserving Data Mining," Proceeding IEEE Second International Advance Computing Conference (IACC), pp. 410-415.

[6] Lomas, N., "Data on 84,000 United Kingdom Prisoners is Lost", Retrieved Sept. 12, 2008, http://news.cnet.com/8301-1009_3-10024550-83.html.

[7] BBC News Brown Apologises for Records Loss, Retrieved September, 12, 2008, http://news.bbc.co.uk/2/hi/uk_news/politics/7104945.stm

[8] Kaplan, D., Hackers Steal 22,000 Social Security Numbers from University of Missouri Database, Retrieved September 2008, http://www.scmagazineus.com/Hackers-steal-22000-Social-Security-numbers-from-Univ.-of-Missouri-database/article/34964/

[9] Goodin, D., "Hackers Infiltrate TD Ameritrade Client Database", Retrieved September 2008, http://www.channelregister.co.uk/2007/09/15/ameritrade_database_burgled/

[10] Liu, L.; Kantarcioglu, M., and Thuraisingham, B., (2009). "Privacy Preserving Decision Tree Mining from Perturbed Data", Proceedings 42nd Hawaii International Conference System Sciences (HICSS '09)

[11] Zhu, Y.; Huang,; Yang, W.; Li, D.; Luo, Y., and F. Dong, (2009). "Three New Approaches to Privacy-Preserving Add to Multiply Protocol and Its Application", Proceeding Second International Workshop Knowledge Discovery and Data Mining, (WKDD '09), pp. 554-558.

[12] Vaidya, J.; and Clifton, C., (2002). "Privacy Preserving Association Rule Mining in Vertically Partitioned Data", Proceeding Eighth ACM SIGKDD International Conference Knowledge Discovery and Data Mining (KDD '02), July, pp. 23- 26.

[13] Shaneck, M.; and Kim, Y., (2010). "Efficient Cryptographic Primitives for Private Data Mining", Proceeding 43rd Hawaii International Conference System Sciences (HICSS), pp. 1-9.

[14] Aggarwal, C., and Yu, P., (2008). "Privacy - Preserving Data Mining:, Models and Algorithms", Springer.

[15] Sweeney, L., (2002). "k-Anonymity: A Model for Protecting Privacy", International Journal Uncertainty, Fuzziness and Knowledge-based Systems, May, Volume. 10, pp. 557-570.

[16] Dowd, J.; Xu, S., and Zhang, W., (2006). "Privacy-Preserving Decision Tree Mining Based on Random Substitions", Proceeding International Conference Emerging Trends in Information and Communication Security (ETRICS '06), pp. 145-159.

[17] Bu, S.; Lakshmanan, L.; Ng, R.; and Ramesh, G., (2007). "Preservation of Patterns and Input-Output Privacy", cedinging IEEE 23rd International Conference Data Eng., April, pp. 696-705.

[18] Russell, S., and Peter, (2002). Artificial Intelligence: A Modern Approach 2/ E. Prentice-Hall.

[19] Fong, P. K., (2008). "Privacy Preservation for Training Data Sets in Database: Application to Decision Tree Learning", Master's Thesis, Department of Computer Science, University of Victoria.

*****

HANDLING SELFISHNESS USING WATCHDOG TECHNIQUES

Smrithy A. S.[10] C. Padmavathy[11]

ABSTRACT

MOBILE ad hoc networks (MANETs) have attracted a lot of attention due to the popularity of mobile devices and the advances in wireless communication technologies. A MANET is a peer-to-peer multi hop mobile wireless network that has neither a fixed infrastructure nor a central server. Each node in a MANET acts as a router, and communicates with each other. The resource and mobility constraints of mobile nodes may lead to network partitioning or performance degradation. Several data replication techniques have been proposed to minimize performance degradation. Most of them assume that all mobile nodes collaborate fully in terms of sharing their memory space. In reality, however, some nodes may selfishly decide only to cooperate partially, or not at all, with other nodes. These selfish nodes could then reduce the overall data accessibility in the network. Due to such problem the overall process of MANET got affected. In this work, we examine the impact of selfish nodes in a mobile ad hoc network from the perspective of replica allocation. We term this selfish replica allocation. Our work was motivated by the fact that a selfish replica allocation could lead to overall poor data accessibility in a MANET. The proposed strategies are inspired by real-world observations in economics in terms of credit risk and in human friendship management in terms of choosing one's friends completely at one's own discretion. We applied the notion of credit risk from economics to detect selfish nodes. Every node in a MANET calculates credit risk information on other connected nodes individually to measure the degree of selfishness.

Our proposed system includes Watchdogs, which are used to detect selfish nodes in computer networks and novel replica allocation techniques to handle the selfish replica allocation appropriately. This is a collaborative approach to reduce the detection time and to improve the accuracy of watchdogs. Numerical results show that our collaborative watchdog can dramatically reduce the overall detection time with a reduced overhead.

KEYWORDS

MANET, ad hoc, Wired Networks etc.

I. INTRODUCTION

Mobile Ad Hoc Networking (MANET) provides a means of wireless routing; two wireless nodes, out of range, wishing to communicate, can leverage nodes in between to carry packets. This is accomplished by ad hoc routing, the mechanism layered over a network providing a route discovery and recovery mechanism allowing for data to be transported from node to node. There are multiple classifications of ad hoc routing: on demand, pre defined / table driven, geo-location aware, power aware, hybrid, etc Considering an on demand protocol, when a node must send data to another, if an entry for the destination does not exist in the sender's routing tables, it will broadcast a route request (RREQ) message to propagate until a route is found. A returned route reply (RREP) message communicates the path back to the originating node. In addition, if a route is lost because nodes moved or dropped from the network, a route error (RERR) will propagate back to the sender and a route discovery will be repeated. Utilizing a discovered route, the packet can then be routed from source to destination with the cooperation of chosen nodes or paths. This extra layer on top of a wireless system presents potential security issues that can disrupt effective communication

Data Replication in MANET

In mobile ad hoc networks (MANETs), since mobile nodes move freely, network partition may occur, where nodes in one partition cannot access data held by nodes in other partitions. Thus, data availability (i.e., the number of successful data accesses over the total number of data accesses) in MANETs is lower than that in conventional wired networks. Data replication has been widely used to improve data availability in distributed systems, and we will apply this technique to MANETs. By replicating data at mobile nodes which are not the owners of the original data, data availability can be improved because there are multiple replicas in the network and the probability of finding one copy of the data is higher. Also, data replication can reduce the query delay since mobile nodes can obtain the data from some nearby replicas. However, most mobile nodes only have limited storage space, bandwidth, and power, and hence it is impossible for one node to collect and hold all the data considering these constraints.

In ad hoc networks, since mobile hosts move freely, disconnections occur frequently, and this causes frequent network partition. If a network is partitioned into two networks due to the migrations of mobile hosts, mobile hosts in one of the partitions cannot access data items held by mobile hosts in the other. Thus, data accessibility in ad hoc networks is lower than that in conventional fixed networks. In ad hoc networks, it is very important to prevent the deterioration of data accessibility at the point of network partition. A possible and promising solution is the replication of data items at mobile hosts which are not the owners of the original data. Since mobile hosts generally have poor resources, it is usually impossible for them to have replicas of all data items in the network. A node may act selfishly, i.e., using its limited resource only for its own benefit, since each node in a MANET has resource constraints, such as battery and storage limitations. A node would like to enjoy the benefits provided by the resources of other nodes, but it may not make its own resource available to help others. Such selfish behavior can potentially lead to a wide range of problems for a MANET. Existing research on selfish behaviors in a MANET mostly focus on network issues. For example, selfish nodes may not transmit data to others to conserve their own batteries. Although network issues are important in a MANET, replica allocation is also crucial, since the ultimate goal of using a MANET is to provide data services to users.

Network partitions can occur frequently, since nodes move freely in a MANET, causing some data to be often inaccessible to some of the nodes. Hence, data accessibility is often an

[10] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, smrithya16@gmail.com
[11] Associate Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, pdrudhra.2007@gmail.com

important performance metric in a MANET. Data are usually replicated at nodes, other than the original owners, to increase data accessibility to cope with frequent network partitions. Data replication can simultaneously improve data accessibility and reduce query delay in MANET; the nodes have sufficient memory space to hold both all the replicas and the original data. A node may act selfishly by using its limited resource only for its own benefit, since each node in a MANET has resource constraints, such as battery and storage limitations. A node would like to enjoy the benefits provided by the resources of other nodes, but it may not make its own resource available to help others. Such selfish behavior can potentially lead to a wide range of problems for a MANET. To solve such problems we examine the impact of selfish nodes in a mobile ad hoc network from the perspective of replica allocation. We term this selfish replica allocation. In particular, we develop a selfish node detection algorithm that considers partial selfishness and novel replica allocation techniques to properly cope with selfish replica allocation.Our objective of this project is to address the problem of selfishness in the context of replica allocation in a MANET, i.e., a selfish node may not share its own memory space to store replica for the benefit of other nodes. First, we examine the impact of selfish nodes in a mobile ad hoc network from the perspective of replica allocation. We term this selfish replica allocation. Second, we develop a selfish node detection algorithm that considers partial selfishness and novel replica allocation techniques to properly cope with selfish replica allocation. We consider that watchdogs are the appropriate mechanism to detect these situations. Essentially, watchdog systems overhear wireless traffic and analyze it to decide if neighbouring nodes are not cooperating. Several works have studied the impact of node selfishness on MANETs proposing different detection mechanisms. In a bayesian watchdog was introduced, as a way to improve the accuracy of the detection.

Our scope of the project is to address the problem of selfish nodes from the replica allocation perspective. We term this problem selfish replica allocation. Our work was motivated by the fact that a selfish replica allocation could lead to overall poor data accessibility in a MANET. We have proposed a selfish node detection method and novel replica allocation techniques to handle the selfish replica allocation appropriately. The proposed strategies collaborative watch dog is inspired by the real-world observations in economics in terms of credit risk and in human friendship management in terms of choosing one's friends completely at one's own discretion. We applied the notion of credit risk from economics to detect selfish nodes. Every node in a MANET calculates credit risk information on other connected nodes individually to measure the degree of selfishness. Since traditional replica allocation techniques failed to consider selfish nodes, we also proposed novel replica allocation techniques. Extensive simulation shows that the proposed strategies outperform existing representative cooperative replica allocation techniques in terms of data accessibility, communication cost, and query delay.

II. EXISTING SYSTEM

Multi-hop communication in mobile ad hoc networks (MANETs) requires collaboration among nodes, which forward packets for one another. Most studies of ad hoc networks assume that nodes can be programmed to always perform this forwarding functionality. In commercial deployment of MANETs, however, some nodes may refuse to forward packets in order to conserve their limited resources (for example, energy), resulting in traffic disruption. Nodes exhibiting such behavior are termed selfish. Selfishness is usually passive behavior. Additionally, malicious nodes may intentionally, and without concern about their own resources, attempt to disrupt network operations by mounting denial-of-service attacks or by actively degrading the network performance. For example, malicious nodes could disrupt routing operation by advertising non-existent routes or sub-optimal routes. Selfish and malicious behaviors are usually distinguished based on the node's intent. Network disruption is a side effect of the behavior of a selfish node, while disrupting the network is the intent of malicious nodes. One way to recognize and isolate such disruptive node behavior is through trust management mechanisms.

Various techniques have been proposed to handle the problem of selfish behavior from the network perspective. As described, the techniques handling selfish nodes can be classified into three categories: reputation-based, credit-payment, and game theory-based techniques. In reputation-based a large number of schemes belong to the first category, with varying implementations. One advantage of such schemes could be their quick convergence in detecting node misbehavior, especially in a large ad hoc network, due to increased information regarding a particular node's behavior. However, this approach has two potential drawbacks: they often assume that nodes that send reputation information about their peers are themselves trustworthy; and they are subject to collusion among nodes that misreport reputation information. In credit-payment techniques, each node gives a credit to others, as a reward for data forwarding. The acquired credit is then used to send data to others. The game theory-based techniques assume that all rational nodes can determine their own optimal strategies to maximize their profit. The game theory-based techniques want to find the Nash Equilibrium point to maximize system performance.

III. PROPOSED SYSTEM

We have proposed a selfish node detection method and novel replica allocation techniques to handle the selfish replica allocation appropriately. The proposed strategies are inspired by the real-world observations in economics in terms of credit risk and in human friendship management in terms of choosing one's friends completely at one's own discretion. We applied the notion of credit risk from economics to detect selfish nodes. Every node in a MANET calculates credit risk information on other connected nodes individually to measure the degree of selfishness. Since traditional replica allocation techniques failed to consider selfish nodes, we also proposed novel replica allocation techniques.

First we detect the selfish node by self replica allocation. Use those replica we devise novel replica allocation techniques with the developed selfish node detection method. They are based on the concept of a self-centered friendship tree (SCF-tree) and its variation to achieve high data accessibility with low communication cost in the presence of selfish nodes. The SCF-tree is inspired by our human friendship management in the real world. In the real world, a friendship, which is a form of social bond, is made individually. For example, although 'A' and 'B' are friends, the friends of A are not always the same as the friends of B. With the help of SCF tree, we aim to reduce the communication cost, while still achieving good data accessibility. Also, Watchdogs are used to detect selfish nodes in computer networks. A way to reduce the detection time and to improve the accuracy of watchdogs is the collaborative

approach. A collaborative watchdog based on contact dissemination of the detected selfish nodes is applied. An analytical model to evaluate the detection time and the cost of this collaborative approach is introduced. The technical contributions of this paper can be summarized as follows:

- ***Recognizing the selfish replica allocation problem***
  We view a selfish node in a MANET from the perspective of data replication, and recognize that selfish replica allocation can lead to degraded data accessibility in a MANET.

- ***Detecting the fully or Partially selfish Nodes Effectively***
  We devise a selfish node detection method that can measure the degree of selfishness using watchdog.

- ***Allocating Replica Effectively***
  We propose a set of replica allocation techniques that use the self-centered friendship tree to reduce communication cost, while achieving good data accessibility.

- ***Verifying the Proposed Strategy***
  The simulation results verify the efficacy of our proposed strategy.

After building the SCF-tree, a node allocates replica at every relocation period. Each node asks non selfish nodes within its SCF-tree to hold replica when it cannot hold replica in its local memory space. Since the SCF-tree based replica allocation is performed in a fully distributed manner, each node determines replica allocation individually without any communication with other nodes. Also, watchdog reduces the detection time of selfish nodes based on contact dissemination. If one node has previously detected a selfish node using its watchdog it can spread this information to other nodes when a contact occurs.

We say that a node has a positive if it knows the selfish node. The detection of contacts between nodes is straightforward using the node's watchdog. Notice that the watchdog is overhearing the packets of the neighborhood; thus, when it starts receiving packets from a new node it is assumed to be a new contact. Then, the node transmits one message including all known positives it knows to this new contacted node. The number of messages needed for this task is the overhead of the collaborative watchdog.the advantages of proposed system is:

- Cooperative replica allocation techniques in terms of data accessibility, communication cost, and query delay so it is efficient.
- A collaborative watchdog can reduce the overall detection time with a reduced cost in term of message overhead. The network is modeled as a set of N wireless mobile nodes with C collaborative nodes and S selfish nodes (N = C +S). At a specific period, or relocation period, each node executes the following procedures:
  - Each node detects the selfish nodes based on credit risk scores (CR).
  - Each node makes its own (partial) topology graph and builds its own SCF-tree by excluding selfish nodes.
  - Based on SCF-tree, each node allocates replica in a fully distributed manner.

The CR score is updated accordingly during the query processing phase to effectively measure the "degree of selfishness".

$$\text{Credit Risk} = \frac{expected\ risk}{expected\ value}.$$

A node wants to know if another node is believable, in the sense that a replica can be paid back, or served upon request to share a memory space in a MANET. With the measured degree of selfishness, a novel tree that represents relationships among nodes in a MANET is proposed for replica allocation termed the SCF-tree. The key strength of the SCF-tree-based replica allocation techniques is that it can minimize the communication cost, while achieving high data accessibility. This is because each node detects selfishness and makes replica allocation at its own discretion, without forming any group or engaging in lengthy negotiations.

At each relocation period, node Ni detects selfish nodes based on nCRki. Each node may have its own initial value of Pki as a system parameter. Interestingly, the initial value of Pki can represent the basic attitude toward strangers. For instance, if the initial value equals zero, node Ni always treats a new node as a non selfish node. Therefore, Ni can cooperate with strangers easily for cooperative replica sharing. Replicas of data items are allocated by allocation techniques. After replica allocation, Ni sets NDki and SSki accordingly. Recall that both NDki and SSki are estimated values, not accurate ones. The estimated values are adjusted at query processing time, according to Algorithm 2.

After building the SCF-tree, a node allocates replica at every relocation period. Each node asks non selfish nodes within its SCF-tree to hold replica when it cannot hold replica in its local memory space. Since the SCF-tree based replica allocation is performed in a fully distributed manner, each node determines replica allocation individually without any communication with other nodes.

In this module, detection time of selfish nodes have to be reduced based on contact dissemination. If one node has previously detected a selfish node using its watchdog it can spread this information to other nodes when a contact occurs. We say that a node has a positive if it knows the selfish node. The watchdog is overhearing the packets of the neighborhood. Thus, when it starts receiving packets from a new node it is assumed to be a new contact. Then, the node transmits one message including all known positives it knows to this new contacted node. The number of messages needed for this task is the overhead of the collaborative watchdog.

Formally, we have a network of N wireless mobile nodes, with C collaborative nodes and S selfish nodes. Initially, the collaborative nodes have no information about the selfish nodes. A collaborative node can have a positive when a contact occurs in the following way:

> **Selfish Contact:** one of the nodes is the selfish node. Then, the collaborative node can detect it using its watchdog and have a positive about this selfish node. Nevertheless, a contact does not always imply detection. To model this fact, we introduce a probability of detection (pd). This probability depends on the effectiveness of the watchdog and the type of contact (for example if the contact time is very low, the watchdog does not have enough information to evaluate if the node is selfish or not).

Collaborative Contact: both nodes are collaborative. Then, if one of them has one or more positives, it can transmit this information to the other node; so, from that moment, both nodes have these positives. As in the selfish contact case, a contact does not always imply collaboration. We model this with the probability of collaboration (pc). The degree of collaboration is a global parameter of the network to be evaluated. This value is used to reflect that either a message with the information about the selfish nodes is lost or that a node temporally does not collaborate (for example, due to a failure or simply because it is switched off). In real networks, full collaboration (pc = 1) is almost impossible.

Although defining a reaction scheme is out of the scope of this paper, there are basically two approaches in the literature: isolation and incentivation. Isolation methods are intended to keep the misbehaving nodes outside the network, excluding them from all kinds of communication. Incentivation methods try to convince the selfish nodes to change their behaviour, and become collaborative instead of selfish, using a virtual payment scheme or a similar mechanism.

A node has 2 states: NOINFO, when the node has no information about the selfish node, and POSITIVE when the node knows who the selfish node. All nodes have an initial state of NOINFO and they can change their initial state when a contact occurs. Using a contact rate λ we can model the network using a Continuous Time Markov Chain (CTMC) with states si =(c), where c represents the number of collaborative nodes in the POSITIVE state. At the beginning, all nodes are in NOINFO state. Then, when a contact occurs, c can increase by one.

First, we derive a basic model for S = 1. In this case, a collaborative node has 2 states: NOINFO, when the node has no information about the selfish node, and POSITIVE when the node knows who the selfish node is (it has a positive). All nodes have an initial state of NOINFO and they can change their initial state when a contact occurs. Using a contact rate λ we can model the network using a Continuous Time Markov Chain (CTMC) with states si = (c), where c represents the number of collaborative nodes in the POSITIVE state. At the beginning, all nodes are in NOINFO state. Then, when a contact occurs, c can increase by one. The final (absorbing) state is when c = C. So, this can be modeled using a CTMC with an initial state s1 = (0), τ = C transient states, and one (υ = 1) absorbing state sτ+1 = (C + 1). Then, the transition matrix P in canonical form is:

$$\mathbf{P} = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

Where I is a υ × υ identity matrix (in this case 1), 0 is a υ × τ zero matrix, Q is a τ × τ matrix with elements pij denoting the transition rate from transient state si to transient state sj and R is a τ × υ matrix with elements pij denoting the transition rate from transient state si to the absorbing state sj. Since we only need the expected time from state s1 = (0) to absorption, the detection time Td, is:

$$T_d = \mathbf{v_1 N v}$$

Where v1 = [1, 0, . . . , 0].

Assume both nodes are collaborative. Then, if one of them has one or more positives, it can transmit this information to the other node; so, from that moment, both nodes have these positives. We model this with the probability of collaboration (pc). The degree of collaboration is a global parameter of the network to be evaluated. This value is used to reflect that either a message with the information about the selfish nodes is lost or that a node temporally does not collaborate.

IV. CONCLUSIONS

In contrast to the network viewpoint, we have addressed the problem of selfish nodes from the replica allocation perspective. We term this problem selfish replica allocation. Our work was motivated by the fact that a selfish replic allocation could lead to overall poor data accessibility in a MANET. We have proposed a selfish node detectionmethod and novel replica allocation techniques to handlethe selfish replica allocation appropriately. We proposed watchdog technique for selfish node dection.

V. REFERENCES

[1] Hara, Takahiro, and Sanjay K. Madria, "Data Replication for Improving Data Accessibility in Ad Hoc Networks".

[2] Ullah Khan, Samee, and Ahmad, Ishfaq, "A Pure Nash Equilibrium - Based Game Theoretical Method for Data Replication across Multiple Servers".

[3] Hara, Takahiro, and Sanjay K. Madria, "Consistency Management Strategies for Data Replication in Mobile Ad Hoc Networks".

[4] Hales, David. "From Selfish Nodes to Cooperative Networks – Emergent Link-based incentives in Peer-to-Peer Networks".

[5] Anderegg, Luzi, and Eidenbenz, Stephan. "Ad hoc-VCG: A Truthful and Cost-Efficient Routing Protocol for Mobile Ad hoc Networks with Selfish Agents".

[6] Byung, Gon Chun; Kamalika Chaudhuri; Hoeteck Weez; Marco Barrenox; Christos H., and Papadimitriouy, John Kubiatowicz, "Selfish Caching in Distributed Systems: A Game Theoretic Analysis".

[7] Cao, Guohong; Liangzhong, Yin, and Chita R. Das, "Cooperative Cache - Based Data Access in Ad Hoc Networks".

[8] Ernesto Damiani, Sabrina De Capitani di Vimercati; Stefano Paraboschi, and Pierangela Samarati, "Managing and Sharing Servents Reputations in P2P Systems".

[9] Balakrishnan, Kashyap; Jing, Deng, and Pramod K. Varshney, "TWOACK: Preventing Selfishness in Mobile Ad Hoc Networks".

[10] Mei, Li; Wang-Chien Leean, and Sivasubramaniam, "Efficient Peer-to-Peer Information Sharing over Mobile Ad Hoc Networks".

*****

RECORD DEDUPLICATION APPROACH USING MODIFIED BAT ALGORITHM

Divya P.[12] Mookhambika N.[13]

ABSTRACT

Record deduplication can be done using many techniques such as genetic programming approach, PSO and bat algorithms under the domain area data mining. Record deduplication is used to identify records in a data repository that refer to the same real world entity or object in spite of misspelling words, typos, different writing styles or even different schema representations or data types.

Existing system aims at providing Unsupervised Duplication Detection method which can be used to identify and remove the duplicate records from different data sources. UDD, which for a given query, can effectively identify duplicates from the query result records of multiple web databases. After removal of the same source duplicates, the "presumed" non duplicate records from the same source can be used as training examples alleviating the burden of users having to manually label training examples. Starting from the non duplicate set, the two cooperating classifiers, a Weighted Component Similarity Summing Classifier (WCSS) and Support Vector Machine (SVM) are used to iteratively identify the duplicate records from the non duplicate record and present a genetic programming (GP) approach to record deduplication. Their approach combines several different pieces of evidence extracted from the data content to produce a deduplication function that is able to identify whether two or more entries in a repository are replicas or not. Since record deduplication is a time consuming task even for small repositories, their aim is to foster a method that finds a proper combination of the best pieces of evidence, thus yielding a deduplication function that maximizes performance using a small representative portion of the corresponding data for training purposes; but the optimization of process is less.

Our proposed system has a new technique modified bat algorithm for record duplication. The motivation behind is to create a resilient and effective method that uses Data Mining algorithms. The system shares many similarities with evolutionary computation techniques such as Genetic programming approach. The system is initialized with a population of random solutions and searches for optima by updating bat generations.

However, unlike G.P. modified bat has no evolution operators such as crossover and mutation. We also compare the proposed algorithm with other existing algorithms, including GP.

KEYWORDS

MANET, ad hoc, Wired Networks etc.

[12] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, divyaparakkil@gmail.com
[13] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

INTRODUCTION

Record deduplication is the task of identifying, in a data repository, records that refer to the same real world entity or object in spite of misspelling words, typos, different writing styles or even different schema representations or data types. In this Research, the existing is a genetic programming (GP) approach to record deduplication. Our EXISTING approach combines several different pieces of evidence extracted from the data content to produce a deduplication function that is able to identify whether two or more entries in a repository are replicas or not. Since record deduplication is a time consuming task even for small repositories, our aim is to foster a method that finds a proper combination of the best pieces of evidence, thus yielding a deduplication function that maximizes performance using a small representative portion of the corresponding data for training purposes. Then, this function can be used on the remaining data or even applied to other repositories with similar characteristics.

Moreover, new additional data can be treated similarly by the suggested function, as long as there are no abrupt changes in the data patterns, something that is very improbable in large data repositories. The existing genetic programming approach considers all the data to found the duplicate records.

Record Deduplication

Deduplication is a key operation in integrating data from multiple sources. The main challenge in this task is designing a function that can resolve when a pair of records refers to the same entity in spite of various data inconsistencies. Record deduplication is the task of identifying, in a data repository, records that refer to the same real world entity or object in spite of misspelling words, typos, different writing styles or even different schema representations or data types. Genetic programming approach to record deduplication is done under the area of domain Datamining.

DATAMINING

Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

Datamining in Record Deduplication

Several systems that rely on consistent data to offer high-quality services, such as digital libraries and e-commerce brokers, may be affected by the existence of duplicates, quasi replicas, or near-duplicate entries in their repositories. Because of that, there have been significant investments from private and government organizations for developing methods for removing replicas from its data repositories. This is due to the fact that clean and replica-free repositories not only allow the

retrieval of higher quality information but also lead to more concise data and to potential savings in computational time and resources to process this data. Deduplication one data set or linking several data sets are increasingly important tasks in the data preparation steps of many data mining projects.

The aim of such linkages is to match all records relating to the same entity. Research interest in this area has increased in recent years, with techniques originating from statistics, machine learning, information retrieval, and database research being combined and applied to improve the linkage quality, as well as to increase performance and efficiency when deduplication or linking very large data sets.

OVERVIEW OF GENETIC PROGRAMMING APPROACH IN RECORD DEDUPLICATION

GP evolves a population of length-free data structures, also called individuals, each one representing a single solution to a given problem. During evolutionary process, the individuals are handled and modified by genetic operations such as reproduction, crossover, and mutation, in an iterative way that is expected to spawn better individuals (solutions to the proposed problem) in the subsequent generations. In this work, the GP evolutionary process is guided by a generational evolutionary algorithm. This means that there are well defined and distinct generation cycles.

We adopted this approach since it captures the basic idea behind several evolutionary algorithms. The steps of this algorithm are the following:

1. Initialize the population (with random or user provided individuals).
2. Evaluate all individuals in the present population, assigning a numeric rating or fitness value to each one.
3. If the termination criterion is fulfilled, then execute the last step. Otherwise continue.
4. Reproduce the best n individuals into the next generation population.
5. Select m individuals that will compose the next generation with the best parents.
6. Apply the genetic operations to all individuals selected. Their offspring will compose the next Population. Replace the existing generation by generated population and go back to Step 2.
7. Present the best individual(s) in the population as the output of the evolutionary process.

Genetic programming approach to record deduplication that combines several different pieces of evidence extracted from the data content to find a deduplication function that is able to identify whether two entries in a repository are replicas or not. As shown by our experiments, our approach outperforms an existing state-of-the-art method found in the literature.

Moreover, the suggested functions are computationally less demanding since they use fewer evidence. In addition, our genetic programming approach is capable of automatically adapting these functions to a given fixed replica identification boundary, freeing the user from the burden of having to choose and tune this parameter.

In our GP experimental environment, individuals are evaluated on how well they learn to predict good answers to a given problem, using the set of functions and terminals available. The resulting value is also called raw fitness and the evaluation functions are called fitness functions.

Each individual is represented by a tree which is a potential solution to the given problem. From that starting point the algorithm iteratively transforms the population into a population with better individuals by applying a number of genetic operators.

These operations are applied to individuals who have been selected based on a fitness measure which determines how close a specific individual is to the desired solution. The three genetic operators typically used in genetic programming are:

- **Reproduction:** An individual is copied without modification.
- **Crossover:** Two selected individuals are recombined into a new individual.
- **Mutation:** A random modification is applied to the selected individual.

The algorithm stops as soon as either the configured maximum number of iterations or a user-defined stop condition is reached.

EXISTING SYSTEM

In existing system they present a genetic programming (GP) approach to record deduplication. Their GP-based approach is also able to automatically find effective deduplication functions, even when the most suitable similarity function for each record attribute is not known in advance. This is extremely useful for the non-specialized user, who does not have to worry about selecting these functions for the deduplication task.

In addition, they show that their approach is also able to adapt the suggested deduplication function to changes on the replica identification boundaries used to classify a pair of records as a match or not. This releases the user from the burden of having to choose and tune these parameter values.

Their approach combines several different pieces of evidence extracted from the data content to produce a deduplication function that is able to identify whether two or more entries in a repository are replicas or not. Since record deduplication is a time consuming task even for small repositories, their aim is to foster a method that finds a proper combination of the best pieces of evidence, thus yielding a deduplication function that maximizes performance using a small representative portion of the corresponding data for training purposes.

The draw backs in the existing system are the optimization of this process is less.Certain optimization problems cannot be solved by means of genetic algorithms. This occurs due to poorly known fitness functions which generate bad chromosome blocks in spite of the fact that only good chromosome blocks cross-over.

- There is no absolute assurance that a genetic algorithm will find a global optimum. It happens very often when the populations have a lot of subjects.

PROPOSED SYSTEM

By idealizing some of the echolocation characteristics of micro-bats, we can develop various bat-inspired algorithms or bat algorithms. Here we developed Modified Bat Algorithm with Doppler Effect. For simplicity, here some of the approximate or idealized rules:

1. All bats use echolocation to sense distance, and they also „know" the difference between food/prey and background barriers in some magical way;

2. Bats fly randomly with velocity vi at position xi with a fixed frequency fmin, varying wavelength λ and loudness A0 to search for prey. They can automatically adjust the wavelength (or frequency) of their emitted pulses and adjust the rate of pulse emission r ∈ [0, 1], depending on the proximity of their target;

3. Doppler Effect is the change in frequency of a wave for an observer moving relative to the source of the wave. The received frequency is higher (compared to the emitted frequency) during the approach, it is identical at the instant of passing by, and it is lower during the recession.

(I) Where *v*s is positive if the source is moving away from the observer, and negative if the source is moving towards the observer.

$$f = \left(\frac{c}{c + v_s}\right) f_0 \quad \text{----(1)}$$

(II) Where the similar convention applies: *v*r is positive if the observer is moving towards the source, and negative if the:

$$f = \left(\frac{c + v_r}{c}\right) f_0 \quad \text{----(1)}$$

(III) Single equation with both the source and receiver moving.

$$f = \left(\frac{c + v_r}{c + v_s}\right) f_0 \quad \text{-----(1)}$$

C is the velocity of waves in the medium;

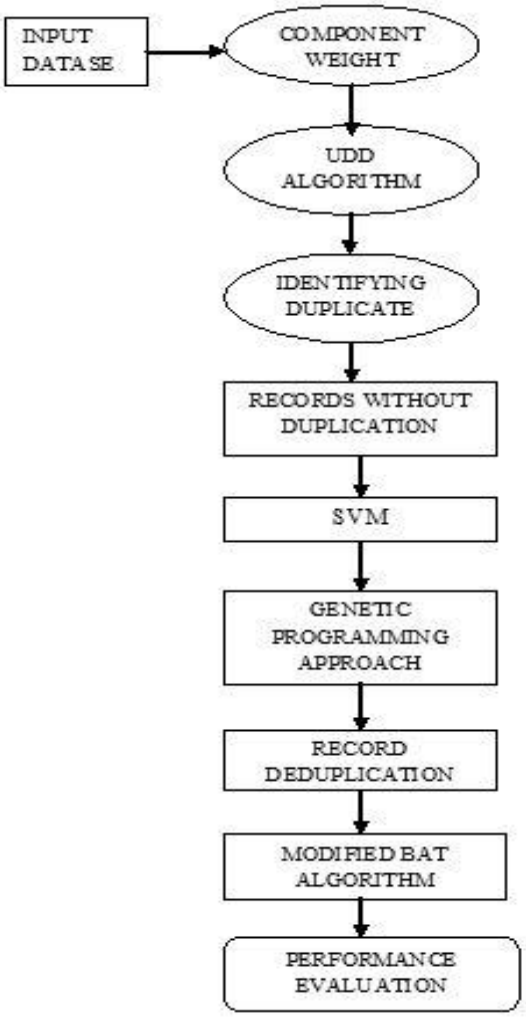
V_r is the velocity of the receiver relative to the medium; if the receiver is moving towards the source;

V_s positive is the velocity of the source relative to the medium; positive if the source is moving away from the receiver.

4. Although the loudness can vary in many ways, we assume that the loudness varies from a large (positive) A0 to a minimum constant value Amin.

Advantages of the proposed system includes that it is easy to implement and there are few parameters to adjust. Compared with GA, all the particles tend to converge to the best solution quickly even in the local version in most cases.

ARCHITECTURE DIAGRAM



Sources: Authors Compilation.

In this figure the user consider the set of documents and collect the all information and keywords from the document. After the preprocessing stage the genetic programming approach is applied to find the duplicate and the non duplicate records. Performed weight assignment for duplicate and non duplicate data and passed to UDD Algorithm which effectively find out the duplicates from similar and dissimilar data sources. Finally it retrieves the non-duplicate records to the user. It finds the records without duplication.

Starting from the non duplicate set, the cooperating classifier, a Weighted Component Similarity Summing Classifier (WCSS) used to iteratively identify the duplicate records from the non duplicate record. To evaluate the similarity between two records, it combines the values of each component in the similarity vector for the two records.

The SVM approach combines several different pieces of evidence extracted from the data content to produce a deduplication function that is able to identify whether two or more entries in a repository are replicas or not.

The present Genetic Programming (GP) approach is to record deduplication. The approach combines several different pieces of evidence extracted from the data content to produce a

deduplication function that is able to identify whether two or more entries in a repository are replicas or not. The algorithm increases the optimization of process and increases the most represented data samples are selected, it finds the best optimization solution to deduplication of the records.

In genetic programming approach is an evolutionary process to increase optimization process, rather than the normal operations performed in the normal record deduplication.

In this step we perform all the genetic operation for record deduplication. After finding the best population or the new population finds the duplicate records in the dataset. To reduce the time complexity of the system performs better than the GP and better optimization solution, unless the genetic it reduces the steps to find the duplicate records.

Bat algorithm performs the better than the genetic programming approach to find the duplicate records and non duplicate records. Bat algorithm reduces the steps and best optimization solution for record deduplication.

ALGORITHM

Implementation of UDD algorithm

Input: Potential duplicate vector set P
Non-duplicate vector set N

Output: Duplicate Vector set D

C_1: a classification algorithm with adjustable parameters W that identifies duplicate vector pairs from P
C_2: a supervised classifier

Algorithm:

1. D=Ø
2. Set the parameters W of C_1 according to N
3. Use C_1 to get a set of duplicate vector pairs d_1 from P
4. Use C_1 to get a set duplicate vector pairs f from N
5. $P=P-d_1$
6. While $|d_1| \neq 0$
7. N'=N-f
8. $D=D+d_1+f$
9. Train C_2 using D and N'
10. Classify p using C_2 and get a set of newly identified duplicate vector pairs d_2
11. $P=P-d_2$
12. $D=D+d_2$
13. Adjust the parameters W of C_1 according to N' and D
14. Use C_1 to get a new set of duplicate vector pairs d_1 from P
15. Use C_1 to get a new set of duplicate vector pairs f from N
16. N=N'

Return D

Implementation of weighted component similarity summing classifier (WCSS)

Input: Duplicate vector set D
Non-duplicate vector set N
Weight scheme co-efficient a

Output: Component Weight W

Algorithm:

1. For i=1 to n
2. $p_i=0$
3. $q_i=0$
4. For each vector $v_k=\{v_{k1},\ldots,v_{kn}\}$in D
5. $P_i=p_i+v_{ki}$
6. $S=\sum^{n}_{i=1} P_i$
7. For i= 1 to n
8. $W_{di}= p_i/S$
9. For each vector $v_k=\{v_{k1},\ldots,v_{kn}\}$in N
10. $q_i=q_i+1-v_{ki}$
11. $S=\sum^{n}_{i=1} q_i$
12. For i= 1 to n
13. $W_{di}= p_i/S$
14. $W_i=a.w_{di}+(1-a)w_{mi}$
15. Return $W=\{w_1,\ldots,w_n\}$

C_1—Weighted Component Similarity Summing (WCSS) Classifier

Implementation of MBAT Algorithm

In the proposed system presents an approach new modified bat algorithm to overcome the difficulty and complexity of the genetic programming Approach .The new algorithm finds the best optimization solution for random selection of the input values and removes the duplicate records in the system. The algorithm reduces the number of the steps in the Genetic programming approach.

Optimization is nothing but selection of a best element from some set of available deternatives. An optimization problem consists of maximizing or minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function.

In the proposed system, we implement the IBAT (Modified Bat) which is a metaheuristic algorithm that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality.

The Modified Bat Algorithm is based on the echolocation behavior of micro-bats with varying pulse rates of emission and loudness with Doppler Effect.

CONCLUSIONS

Duplicate detection is an important step in data integration and this method is based on offline learning techniques, which requires training data. In the Web database scenario, where records to match are greatly query-dependent, a pertained approach is not applicable as the set of records in each query's results is a biased subset of the full data set.

To overcome this problem, it presents an unsupervised, online approach UDD, for detecting duplicates over the query results of multiple Web databases. Two classifiers, WCSS and SVM are used cooperatively in the convergence step of record matching to identify the duplicate pairs from all potential duplicate pairs iteratively.

The genetics programming approach combines several different pieces of evidence extracted from the data content to produce a deduplication function that is able to identify whether two or more entries in a repository are replicas or not. Their aim is to

foster a method that finds a proper combination of the best pieces of evidence, thus yielding a deduplication function that maximizes performance using a small representative portion of the corresponding data for training purposes.

In our proposed system increases the optimization of process and increases the most represented data samples are selected, it finds the best optimization solution to deduplication of the records. MBAT shares many similarities with evolutionary computation techniques such as Genetic Algorithms.

The system is initialized with a population of random solutions and searches for optima by updating generations. MBAT search the best optima by updating generations. In MBAT takes and less error rate when comparing to the GP. It is a one-way information sharing mechanism. The evolution only looks for the best solution. Compared with GP, all modified bat finds the best optima solution for each of the random selection input data.

REFERENCES

[1]. M., Bilenko, and Mooney, R. J., (2003). "Adaptive Duplicate Detection Using Learnable String Similarity Measures", Proceeding 9th ACM SIGKDD International Conference Knowledge Discovery and Data Mining, pp. 39-48.

[2]. Tak-Lam, Wong,*, Wai Lam, (2009). "An Unsupervised Method for Joint Information Extraction and Feature Mining across Different Web Sites", Data & Knowledge Engineering 68, pp. 107–125.

[3]. Donghui, Feng; Gully, Burns, and Eduard, Hovy, (2007). "Extracting Data Records from Unstructured Biomedical Full Text", Proceedings of 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, June, Prague, pp. 837–846.

[4]. Christen, P., (2005). "Probabilistic Data Generation for Deduplication and Data Linkage", Intelligent Data Engineering and Automated Learning, pp. 109-116, Springer.

[5]. Robert, Baumgartner; Oliver, Frolich; Georg Gottlob; Patrick, Harz; Marcus, Herzog, and Peter, Lehmann, (2009). "Web Data Extraction for Business Intelligence: the Lixto Approach".

[6]. Bilenko, M.; R. Mooney; W. Cohen; P. Ravikumar, and S. Fienberg, (2003). "Adaptive Name Matching in Information Integration", IEEE Intelligent Systems, September / October, Volume: 18, Number: 5, pp. 16-23.

[7]. Elmagarmid, A. K.; Ipeirotis, P. G., and Verykios, V. S., (2007). "Duplicate Record Detection: A Survey", IEEE Trans. Knowledge and Data Eng., January, Volume: 19, Number: 1, pp. 1-16.

[8]. Sarawagi, S., and Bhamidipaty, A., (2002). "Interactive Deduplication Using Active Learning", Proceeding 8th ACM SIGKDD International Confernce Knowledge Discovery and Data Mining, pp. 269-278.

[9]. Weifeng, Su; Jiying, Wang, and Frederick H. Lochovsky, (2010). "Record Matching over Query Results from Multiple Web Databases", Knowledge Discovery and Data Mining, April, Volume: 22, Number: 4.

[10]. H. M. de Almeida; M. A. Gonc; Alves, M. Cristo, and P. Calado, (2007). "A Combined Component Approach for Finding Collection-Adapted Ranking Functions Based on Genetic Programming", Proceedings 30th Annual International ACM SIGIR Conference Research and Development in Information Retrieval, pp. 399-406.

[11]. Robert, Burbidge, and Bernard, Buxton, "An Introduction to Support Vector Machines for Data Mining", Computer Science Department, UCL, Gower Street, WC1E 6BT, UK.

[12]. Silva, T.P.C.; de Moura, E. S.; Cavalcanti, J.M.B.; A. S. da Silva; M.G. de Carvalho; and M. A. Gonc¸Alves, (2009). "An Evolutionary Approach for Combining Different Sources of Evidence in Search Engines", Information Systems, Volume: 34, Number: 2, pp. 276-289.

[13]. Raymer, M. L.; Punch, W.F.; Goodman, E. D. and L. A. Kuhn, "Genetic Programming for Improved Data Mining Application to the Biochemistry of Protein Interactions".

[14]. Matthew, Walker, (c2001). "Introductions to Genetic Programming", RoboCup-97: Robot Soccer World Cup I, Springer, October 7, pp. 398-411.

[15]. M. G. de Carvalho; A. H. F. Laender; M. A. Gonc¸Alves, and A. S. da Silva, (2008). "Replica Identification Using Genetic Programming", Proceeding 23rd Annual ACM Symp. Applied Computing (SAC), pp. 1801-1806.

[16]. Bhattacharya, Indrajit, and Getoor, Lise, (2004). "Iterative Record Linkage for Cleaning and Integration", DMKD '04 Proceedings of 9th ACM SIGMOD workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 11-18 ACM New York, NY, USA ©2004.

[17]. Hassan, Rania; Cohanim, Babak, and Olivier de Weck, (2004). "A Copmarison of Particle Swarm Optimization And The Genetic Algorithm**",** Massachusetts Institute of Technology, Cambridge, MA, 02139, 2004.

[18]. Jun Zhang*, Jing-hui Zhong, and Xiao-Min Hu X., (2008). "A Novel Genetic Algorithm with Orthogonal Prediction for Global Numerical Optimization", pp. 31–40, 2008, © Springer-Verlag Berlin Heidelberg.

[19]. Bernstein, Y.; Shokouhi, M., and Zobel, J., (2006). "Compact Features for Detection of Near - Duplicates in Distributed Retrieval", in Proceedings of String Processing and Information Retrieval Symposium (to appear), Glasgow, Schotland.

[20]. Chaudhuri, S.; Ganjam, K.; Ganti, V., and R. Motwani, (2003). "Robust and Efficient Fuzzy Match for Online Data Cleaning", Proceeding ACM SIGMOD International Confernce Management of Data, pp. 313-324.

*****

ACTIVITY-BASED PERSON IDENTIFICATION USING PARTICLE SWARM OPTIMIZATION ALGORITHM

Ambili M. P.[14] D. Malini[15] V. P. Arulkumar[16]

ABSTRACT

A generic view-invariant person identification method has been proposed in existing system, by exploiting the information provided by a multicamera setup and incorporating several activity classes in the person identification procedure. The existing most methods assumed which walk are to be the only activity exploited for person identification, they may perform the other activities different from the walk activity, So we incorporate several activities in order to identify a person.

A multicamera setup, in order to capture the human body from different viewing angles. Fuzzy Vector Quantization (FVQ) and Linear Discriminant Analysis (LDA) are used to analysis the different movement of activity representation. Person identification, activity recognition, and viewing angle specification results are obtained for all the available cameras independently. In proposed system PSO Algorithm is to capture the person identification besides on the walk, run, skip, gallop sideways and jump in place.

Proposed Particle Swarm Optimization (PSO) is a computational method that optimizes problem person identification on several activities by iteratively trying to improve a solution. PSO optimizes a problem by having a population of candidate solutions, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best known position and is also guided toward the best known positions in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions. Experimental results show that the proposed system improves the accuracy, person identification.

KEYWORDS

Fuzzy Vector Quantization (FVQ), Linear Discriminant Analysis, Particle Swarm Optimization etc.

INTRODUCTION

Person identification has been heavily researched in the past few decades due to its importance in security applications. Most of the existing person identification approaches, such as the ones based on face, iris, and fingerprint recognition, assume restricted identification setups and person cooperation. For example, the person should stand at a standard distance in front of a camera and look at a specific point, or have physical contact with sensors. In order to overcome these restrictions, noninvasive biometrics has been exploited, referring to the anatomical or behavioral traits associated with a specific person that can be used for automatic recognition.

Person identification using multiple cameras that observe the person under investigation, while he / she performs several everyday activities. It is evident that the viewing angle plays a significant role in the person identification accuracy. This is due to the fact that the shape of the human body, as well as the body dynamics observed during the execution of activities, differs a lot when the person is observed from different viewing angles. This is the so-called viewing-angle effect.

Most person identification methods proposed in the literature use one camera and assume the same viewing angle during the training and recognition phases. This renders them inappropriate for applications aiming to view-invariant person identification. In order to provide a nonrestrictive person identification setup, the identification accuracy method should ideally not be affected by the viewing angle. Person identification besides on activities of walk, run, skip, gallop sideways, jump jack, jump forward, and jump in place. By properly combining these results, a view - invariant activity - independent person identification method is obtained. To combine all these activities we use Bayesian framework in identification (test) phase it achieves high Identification rates.

Linear Discriminant Analysis (LDA) is used to map the activity representation in a low-dimensional discriminant feature space. In this space, the human activity representation in a test sequence is classified to the nearest class centroid using cosine similarity. This method exploits a robust human body representation. Binary images denoting human body poses can effectively be obtained by using background subtraction techniques.

Activities used for person identification besides "walk" are "run," "skip," "gallop sideways," "jump jack," "jump forward," and "jump in place." It should be noted that experiments have proven that the most distinctive activity was found to be "skip," while "walk" was ranked sixth in discriminant capacity (out of eight activities).In order to overcome the viewing angle effect, three main approaches have been proposed for existing system:

1) Introduction of view-invariant gait representations,
2) Mapping of the gait representation from an arbitrary viewing angle to a reference one and performing identification in this viewing angle,
3) Using multicamera setups.

Activity-based person identification walk is assumed to be one, but not the only, activity, which will be exploited for person identification. We use an uncalibrated multicamera setup, in order to capture the human body from different viewing angles:

1) The use of activities (besides "walk") in person identification.

2) The use of dynemes for person identification. In it has been shown that dynemes can be used for activity recognition. Here, we show that dynemes can also be used for activity-based person identification. Thus,

[14] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, ambilimp13@yahoo.com
[15]Lecturer, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, malinigds@rediffmail.com
[16]Lecturer, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, arulkumarssk@gmail.com

dynemes can be considered as similar human body poses representing similar persons.

3) The use of cumulative fuzzy distances from the dynemes to achieve time-invariant person identity representation.

4) The use of a Bayesian framework to combine the person identification, activity recognition, and viewing angle specification results produced for all the available cameras in the identification phase.

Particle swarm optimization (PSO) is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality.

PSO optimizes a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best known position and is also guided toward the best known positions in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions.

OBJECTIVES OF PROJECT

The main objective of the research is to improve the identification rates by using PSO. Because the existing methods only improve the identification rates based on the classification results. The proposed PSO algorithm solves the optimization problem of the Person identification.

MOTIVATION OF PROJECT

The motivation of the project is to Solves the person identification at different activities in multicamera setup, capturing the activities at the person activities is varying based on person. Several methods solve the person identification, but not give accurate or optimized result, whether the person is correct or not.

BASIC CONCEPTS

Image Processing

Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it.

Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This article is about general techniques that apply to all of them. The *acquisition* of images (producing the input image in the first place) is referred to as imaging.

An image defined in the "real world" is considered to be a function of two real variables, for example, a(x,y) with a as the amplitude (e.g. brightness) of the image at the real coordinate position (x,y).

In a sophisticated image processing system it should be possible to apply specific image processing operations to selected regions. Thus one part of an image (region) might be processed to suppress motion blur while another part might be processed to improve color rendition.

Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. The goal of this manipulation can be divided into three categories: * Image Processing image in -> image out * Image Analysis image in -> measurements out * Image Understanding image in -> high-level description out Image processing is referred to processing of a 2D picture by a computer.

Basic Definitions:

An image defined in the "real world" is considered to be a function of two real variables, for example, a(x, y) with a as the amplitude (e.g. brightness) of the image at the real coordinate position (x, y).

An image may be considered to contain sub-images sometimes referred to as regions-of-interest, ROIs, or simply regions. This concept reflects the fact that images frequently contain collections of objects each of which can be the basis for a region. In a sophisticated image processing system it should be possible to apply specific image processing operations to selected regions. Thus one part of an image (region) might be processed to suppress motion blur while another part might be processed to improve color rendition.

Sequence of Image Processing:

The most requirements for image processing of images is that the images be available in digitized form, that is, arrays of finite length binary words. For digitization, the given Image is sampled on a discrete grid and each sample or pixel is quantized using a finite number of bits. The digitized image is processed by a computer. To display a digital image, it is first converted into analog signal, which is scanned onto a display.

Closely related to image processing are computer graphics and computer vision. In computer graphics, images are manually *made* from physical models of objects, environments, and lighting, instead of being acquired (via imaging devices such as cameras) from *natural* scenes, as in most animated movies. Computer vision, on the other hand, is often considered *high-level* image processing out of which a machine / computer / software intends to decipher the physical contents of an image or a sequence of images (e.g., videos or 3D full-body magnetic resonance scans).

In modern sciences and technologies, images also gain much broader scopes due to the ever growing importance of scientific visualization (of often large-scale complex scientific / experimental data). Examples include microarray data in genetic research, or real-time multi-asset portfolio trading in finance. Before going to processing an image, it is converted into a digital form. Digitization includes sampling of image and quantization of sampled values. After converting the image into bit information, processing is performed. This processing technique may be, Image enhancement, Image reconstruction, and Image compression:

a) Image Enhancement

It refers to accentuation, or sharpening, of image features such as boundaries, or contrast to make a graphic display more useful for display & analysis. This process does not increase the inherent information content in data. It includes gray level & contrast manipulation, noise reduction, edge crispening and sharpening, filtering, interpolation and magnification, pseudo coloring, and so on.

b) Image Restoration

It is concerned with filtering the observed image to minimize the effect of degradations. Effectiveness of image restoration depends on the extent and accuracy of the knowledge of degradation process as well as on filter design. Image restoration differs from image enhancement in that the latter is concerned with more extraction or accentuation of image features.

c) Image Compression

It is concerned with minimizing the number of bits required to represent an image. Application of compression are in broadcast TV, remote sensing via satellite, military communication via aircraft, radar, teleconferencing, facsimile transmission, for educational & business documents, medical images that arise in computer tomography, magnetic resonance imaging and digital radiology, motion, pictures, satellite images, weather maps, geological surveys and so on.

Biometrics

Biometrics is the science and technology of measuring and analyzing biological data. In information technology, biometrics refers to technologies that measure and analyze human body characteristics, such as DNA, fingerprints, eye retinas and irises, voice patterns, facial patterns and hand measurements, for authentication purposes.

Background Subtraction

Background subtraction is a computational vision process of extracting foreground objects in a particular scene. A foreground object can be described as an object of attention which helps in reducing the amount of data to be processed as well as provide important information to the task under consideration.

Often, the foreground object can be thought of as a coherently moving object in a scene. We must emphasize the word coherent here because if a person is walking in front of moving leaves, the person forms the foreground object while leaves though having motion associated with them are considered background due to its repetitive behavior.

In some cases, distance of the moving object also forms a basis for it to be considered a background, e.g if in a scene one person is close to the camera while there is a person far away in background, in this case the nearby person is considered as foreground while the person far away is ignored due to its small size and the lack of information that it provides. Identifying moving objects from a video sequence is a fundamental and critical task in many computer-vision applications. A common approach is to perform background subtraction, which identifies moving objects from the portion of video frame that differs from the background model.

CONCLUSIONS

Unified framework aiming at activity-based person identification proposed, by exploiting the information provided by a multicamera setup and incorporating several activity classes in the person identification procedure.

By applying LDA to this activity video representation, three optimal discriminant subspaces are determined, in order to use the activity video for human identity, activity class, and viewing angle classification. Combining classification results to all the available cameras in the person identification (test) phase using a Bayesian approach, achieves high identification rates.PSO algorithm improves the identification rates by solving the optimization problem. Because the existing system only performs identification based on the classification results.

PSO parameters can have a large impact on optimization performance for person identification rates. Compare to the existing system of the Bayesian approach, proposed PSO Algorithm improves the efficiency person identification rates at each person. PSO-based algorithm was to generate person identification with the less time.

REFERENCES

[1] A., Jain, (2008). "Biometric Recognition: Overview and Recent Advances", Prog. Pattern Recognit. Image Anal., Applicat., Volume: 4756, pp. 13–19.

[2] Wang, J.; She, M.; Nahavandi, S., and Kouzani, A., (2010). "A Review of Vision - Based Gait Recognition Methods for Human Identification", in Proceeding 2010 IEEE International Conference Digital Image Computing: Techniques and Applications, pp. 320–327.

[3] N., Boulgouris; D., Hatzinakos, and K., Plataniotis, (2005). "Gait recognition: A Challenging Signal Processing Technology for Biometric Identification", IEEE Signal Process. Mag., Volume: 22, Number: 6, pp. 78–90, November.

[4] J., Boyd, (2004). "Synchronization of Oscillations for Machine Perception of Gaits", Comput. Vis. Image Understand., Volume: 96, Number: 1, pp. 35–59

[5] M., Nixon, and J., Carter, (2004). "Advances in Automatic Gait Recognition", in Proceeding International Conference Automatic Face and Gesture Recognition, pp. 139–146.

[6] S., Yu D. Tan, and T., Tan, (2006). "Modelling the Effect of View Angle Variation on Appearance-Based Gait Recognition", in Proceeding Computer Vision (ACCV 2006), pp. 807–816.

[7] S., Yu D. Tan, and T., Tan, (2006). "A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition", in Proceeding IEEE 18th International Conference Pattern Recognition, Volume: 4, pp. 441–444.

[8] C., Yam; M., Nixon, and J., Carter, (2001). "Extended Model-Based Automatic Gait Recognition of Walking and

Running", in Audio and Video Based Biometric Person Authentication. New York: Springer, pp. 278–283.

[9] N., Gkalelis; A., Tefas, and I., Pitas, (2010). "Human Identification from Human Movements", in Proceeding 2009 16th IEEE International Conferene Image Processing (ICIP), pp. 2585–2588.

[10] J., Han, and B., Bhanu, (2006). "Individual Recognition Using Gait Energy Image", IEEE Trans. Pattern Anal. Mach. Intell., Volume: 28, Number: 2, pp. 316–322, February.

[11] T., Lam, and R., Lee, (2005). "A New Representation for Human Gait Recognition: Motion Silhouettes Image (MSI)", Adv. Biometrics, Volume: 1, pp. 612–618.

[12] K., Bashir; T., Xiang, and S., Gong, (2010). "Gait Recognition without Subject Cooperation", Pattern Recognit. Lett., Volume: 31, Number: 13, pp. 2052–2060.

[13] D., Tao; X., Li; X., Wu, and S., Maybank, (2007). "General Tensor Discriminant Analysis and Gabor Features for Gait Recognition", IEEE Trans. Pattern Anal. Mach. Intell., October, Volume: 29, Number: 10, pp. 1700–1715.

[14] A., Bobick, and J., Davis, (2001). "The Recognition of Human Movement Using Temporal Templates", IEEE Trans. Pattern Anal. Mach. Intell., Volume: 23, Number: 3, pp. 257–267, March.

[15] Z., Liu, and S., Sarkar, (2006). "Improved Gait Recognition by Gait Dynamics Normalization", IEEE Trans. Pattern Anal. Mach. Intell., Volume: 28, Number: 6, pp. 863–876, Junuary.

[16] N., Boulgouris, and Z., Chi, (2007). "Gait Recognition using Radon Transform and Linear Discriminant Analysis", IEEE Trans. Image Process., Volume: 16, Number: 3, pp. 731–740, March.

[17] F., Jean; R., Bergevin, and A., Albu, (2009). "Computing and Evaluating View Normalized Body Part Trajectories", Image Vis. Comput., Volume: 27, Number: 9, pp. 1272–1284, 2009.

[18] A., Kale; A., Chowdhury, and R., Chellappa, (2003). "Towards a View Invariant Gait Recognition Algorithm", in Proc. IEEE Conference Advanced Video and Signal Based Surveillance, pp. 143–150.

[19] Y., Makihara; R., Sagawa; Y., Mukaigawa; T., Echigo, and Y., Yagi, (2006). "Gait Recognition Using a View Transformation Model in Frequency Domain", in Proceeding Computer Vision (ECCV 2006), pp. 151–163.

[20] W., Kusakunniran; Q., Wu; H., Li, and J., Zhang, (2010). "Multiple Views Gait Recognition Using View Transformation Model Based on Optimized Gait Energy Image", in Proceeding IEEE 12th International Conference Computer Vision Workshops (ICCV Workshops), pp. 1058–1064.

[21] R., Bodor; A., Drenner; D., Fehr; O., Masoud, and N., (2009). Papanikolopoulos, "View-Independent Human Motion Classification Using Image-Based Reconstruction", Image Vis. Comput., Volume: 27, Number: 8, pp. 1194–1206.

[22] G., Shakhnarovich; L., Lee and T., Darrell, (2001). "Integrated Face and Gait Recognition from Multiple views", in Proceeding IEEE Computer Society Conference Computer Vision and Pattern Recognition, Kauai, HI, 2001, Volume; 1, pp. 439–446, Citeseer.

[23] N., Gkalelis; A., Tefas, and I., Pitas, (2008). "Combining Fuzzy Vector Quantization with Linear Discriminant Analysis for Continuous Human Movement Recognition", IEEE Trans. Circuits Syst. Video Technol., Volume; 18, Number: 11, pp. 1511–1521, November.

[24] A., Iosifidis; N., Nikolaidis, and I., Pitas, "Movement Recognition Exploiting Multi-View Information", in Proceedings IEEE International Workshop on Multimedia Signal Processing (MMSP), pp. 427–431.

[25] N., Gkalelis; H., Kim; A., Hilton; N., Nikolaidis, and I., Pitas, (2009). "The i3DPost Multi-View and 3d Human Action / Interaction Database", in Proceeding 6th Conference Visual Media Production, November, pp. 159–168.

[26] Y., Benezeth; P., Jodoin; B., Emile; H., Laurent, and C., Rosenberger, (2009). "Review and Evaluation of Commonly-Implemented Background Subtraction Algorithms", in Proceeding 19th IEEE International Conference Pattern Recognition, pp. 1–4.

*****

FOR ANY CLARIFICATION OR SUGGESTION, WRITE US:

Editor-In-Chief
Pezzottaite Journals,
24, Saraswati Lane, Bohri,
Near Modern Dewan Beverages,
Jammu Tawi – 180002,
Jammu and Kashmir, India.
(Mobile): +91-09419216270 – 71
editorinchief@pezzottaitejournals.net
contactus@pezzottaitejournals.net

REDUCING THE NUMBER OF TRANSMISSIONS USING HYBRID ALGORITHM

Sanitha P. Chandran[17] C. Padmavathy[18]

ABSTRACT

Routing in wireless multihop networks is challenging due to high loss rate and dynamic quality of wireless links. There are two main approaches, static and dynamic, to broadcast algorithms in wireless ad hoc networks. In the static approach, local algorithms determine the status (forwarding / nonforwarding) of each node proactively based on local topology information and a globally known priority function. The local broadcast algorithms based on the static approach cannot guarantee a small sized CDS if the position information is not available; but it is showed that constant approximation factor is achievable using position information. In the dynamic approach, local algorithms determine the status of each node "on-the-fly" based on local topology information and broadcast state information.

Using the dynamic approach, it was recently shown that local broadcast algorithms can achieve a constant approximation factor to the optimum solution when (approximate) position information is available. However, using position information can simplify the problem. Also, in some applications it may not be practical to have position information. Therefore, we wish to know whether local broadcast algorithms based on the dynamic approach can achieve a constant approximation factor without using position information. Using the dynamic approach, the local broadcast algorithms based on the dynamic approach do not require position information to guarantee a constant approximation factor.

Although the existing algorithm can achieve both full delivery and a constant approximation to the optimum solution it underutilizes available bandwidth resources using single transmission rate over the entire network whose links perform work at a lower rate and the network may become disconnected at a higher bit rate.

The proposed system provides an optimal polynomial - time solution to this problem and incorporates the multirate capability inherent in 802.11 net-works into path routing for efficient network performance with guaranteed delivery.

KEYWORDS

Wireless, Network, MANET, Broadcast, Algorithum, Security, Privacy, Routing, Protocols etc.

INTRODUCTION

A mobile ad hoc network (MANET) is an autonomous collection of mobile users (nodes) that A mobile ad hoc network (MANET) is an autonomous collection of mobile users (nodes) that communicate over wireless links. Due to nodal mobility, the network topology may change rapidly and unpredictably over time. The network is decentralized, where network organization and message delivery must be executed by the nodes themselves. MANETs have a wide range of applications from military networks to emergency preparedness telecommunications.

Regardless of the application, there are certain critical features that can determine the efficiency and effectiveness of an ad-hoc network. These features can be categorized into quantitative features and qualitative features. Quantitative critical features includes network settling time, network join, network depart, network recovery time, frequency of updates, memory byte requirement, network scalability number. Qualitative critical features includes knowledge of nodal locations, effect to topology changes, power consciousness, single or multichannel, real time voice services, real time video services.

The existing local broadcast algorithms can be classified based on whether the forwarding nodes are determined statically based on only local topology information or dynamically based on both local topology and broadcast state information. In the static approach, the distinguishing feature of local algorithms over other broadcast algorithms is that using local algorithms any local topology changes can affect only the status of those nodes in the vicinity. Therefore local algorithms can provide scalability as the constructed CDS can be updated, efficiently.

The existing local algorithms in this category typically use a priority function known by all nodes in order to determine the status of each node. The local broadcast algorithms based on the static approach are not able to guarantee a good approximation factor to the optimum solution (i.e., MCDS) but it achieve interesting results such as a constant approximation factor and shortest path preservation if the nodes are provided with position information. In the dynamic approach, the status of each node is determined "on-the-fly" during the broadcast progress.

Using this approach, the constructed CDS may vary from one broadcast instance to another even when the whole network topology and the source node remain unchanged. Consequently, the broadcast algorithms based on the dynamic approach typically have small maintenance cost and are expected to be robust against node failures and network topology changes. Many local broadcast algorithms in this category use local neighbour information to reduce the total number of transmissions and to guarantee full delivery.

The local broadcast algorithm based on 2-hop neighbor information and proves that it guarantees a constant approximation to the optimum solution. The proposed algorithm is both neighbor-designating and selfpruning, i.e., the status of each node is determined by itself and/or other nodes.

In particular using our proposed algorithm, each broadcasting node selects at most one of its neighbors to forward the message. If a node is not selected to forward, it has to decide, on its own, whether or not to forward the message. This can be achieved in two rounds of information exchange. In the first round, each node broadcasts its id to its 1-hop neighbors

[17] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, anithapchandran@gmail.com
[18] Associate Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, pdrudhra.2007@gmail.com

(simply called neighbors). Thus, at the end of the first round, each node has a list of its neighbors. In the second round, each node transmits its id together with the list of its neighbors. The existing broadcast algorithm is a hybrid algorithm; hence every node that broadcasts the message may select some of its neighbors to forward the message. In our proposed broadcast algorithm, every broadcasting node selects at most one of its neighbors. A node has to broadcast the message if it is selected to forward. Other nodes that are not selected have to decide whether or not to broadcast on their own. This decision is made based on self-pruning condition called the coverage condition.

Disadvantages of Existing System

- Broadcast algorithm guarantees that the total number of transmissions is always within a constant factor of the minimum number of required ones. However, the number of transmissions may be further reduced by slightly modifying the broadcast algorithm.
- The local broadcast algorithms based on the static approach cannot guarantee a small sized CDS if the position information is not available.
- The higher bit rates have a shorter transmission range which reduces the network connectivity. As the rate increases, links becomes lossier and the network eventually gets disconnected.

The existing system worked on routing focused only on a single bit rate. Such an assumption, however, considerably under-utilizes available bandwidth resources. Some hyperlinks may be able to sustain a higher transmission rate, while others may only work at a lower rate. Additionally, the transmission range and therefore the network topology change with the rate, which makes multirate a challenge for routing.

Here provide **an optimal polynomial-time solution to this problem** and incorporate the multirate capability inherent in 802.11 net-works into path routing. For each destination, a node then keeps both a forwarding set and a rate used to reach this set. As a result, every two nodes will be connected through a mesh composed of multiple paths, with each node transmitting at a selected rate. The EATX metric described was originally designed considering that nodes transmit at a single bit rate. To account for multiple bit rates, we introduce the expected any-path transmission time (EATT) metric. For EATT, the hyperlink cost for each rate is defined as:

$$d_{ij}^{(r)} = \frac{1}{p_{ij}^{(r)}} \times \frac{s}{r}$$

Where $p_{ij}^{(r)}$ is the hyperlink delivery ratio, S is maximum packet size, and R is the bit rate. The cost $d_{ij}^{(r)}$ is the time taken to transmit a packet of size at a bit rate over a lossy hyperlink with delivery ratio $p_{ij}^{(r)}$. The EATT metric is a generalization of the expected transmission time (ETT) metric commonly used in single-path wireless routing. Note that for each bit rate $r \in R$ we have a different delivery ratio $p_{ij}^{(r)}$ which usually decreases for higher rates. This behavior imposes a tradeoff: A higher bit rate decreases the time of a single packet transmission.

Advantages of Proposed System

- The overall cost of every node to a particular destination is minimized using both forwarding set and a transmission rate for every node.
- The proposed approach is optimal even if packet losses at different receivers are correlated.

Network Model

The network consists of a set of nodes. Every node has a unique id and every packet is stamped by the id of its source node. To model the network, we assume two different nodes u and v are connected by an edge if and only if both nodes are within transmission range R. The distance between the nodes is measured in terms of euclidean distance. Thus the network is connected and static during the broadcast algorithm which guarantees reliable delivery.

Static Broadcasting Using Position Information

The position information reduces total number of broadcasting nodes. In local broadcast algorithms based on the static approach, the status (forwarding / non forwarding) of each node is a function of node's id, fixed priority function and local topology information.

The priority of a node can be determined by its id, degree (number of its 1-hop neighbours) or by its neighbour connectivity ratio. The status functions guarantee in constructing a CDS and also ensure that the constructed CDS is small size within a constant factor of the optimum. The change in local topology affects the status of the nodes in the neighbourhood.

Constant approximation factor is achievable using position information. Thus partition the network area into square cells using geographical information and present simple local algorithms that select a constant number of nodes in each cell. To save energy at each time, algorithm selects one node in each cell as active and puts other nodes in cell to sleep.

The set of active nodes must form a CDS in order to maintain network connectivity and coverage. An important application of constructing a CDS is to employ it as a backbone for routing. When a CDS is constructed the nodes in the set are required to forward packets toward the destination.

Dynamic Broadcasting using self-pruning algorithm

Dynamic broadcast algorithm guarantees full delivery as well as a constant approximation to the optimum solution irrespective of the forwarding node selection criteria and random delay of broadcasting nodes. Using the dynamic approach, the status (forwarding/ non forwarding) of each node is determined "on-the-fly" as the broadcasting message propagates in the network.

In neighbour-designating algorithm each forwarding node selects a subset of its neighbours to forward the packet and in self-pruning algorithms each node determines its own status based on a self-pruning condition after receiving the first or several copies of the message. Self-pruning broadcast algorithms are able to guarantee both full delivery and a

constant approximation factor to the optimum solution (MCDS).

However this algorithm uses position information to design a strong self-pruning condition. Each node has a list of its 2-hop neighbours. This can be achieved in two rounds of information exchange. In the first round, each node broadcasts its id to its 1-hop neighbours. Thus at the end of the first round, each node has a list of its neighbours.

In the second round, each node transmits its id together with the list of its neighbours. In our proposed broadcast algorithm, every broadcasting node selects at most one of its neighbours. A node has to broadcast the message if it is selected to forward. Other nodes that are not selected have to decide whether or not to broadcast on their own. This decision is made based on a self-pruning condition called coverage condition.

Hybrid Broadcast Algorithm

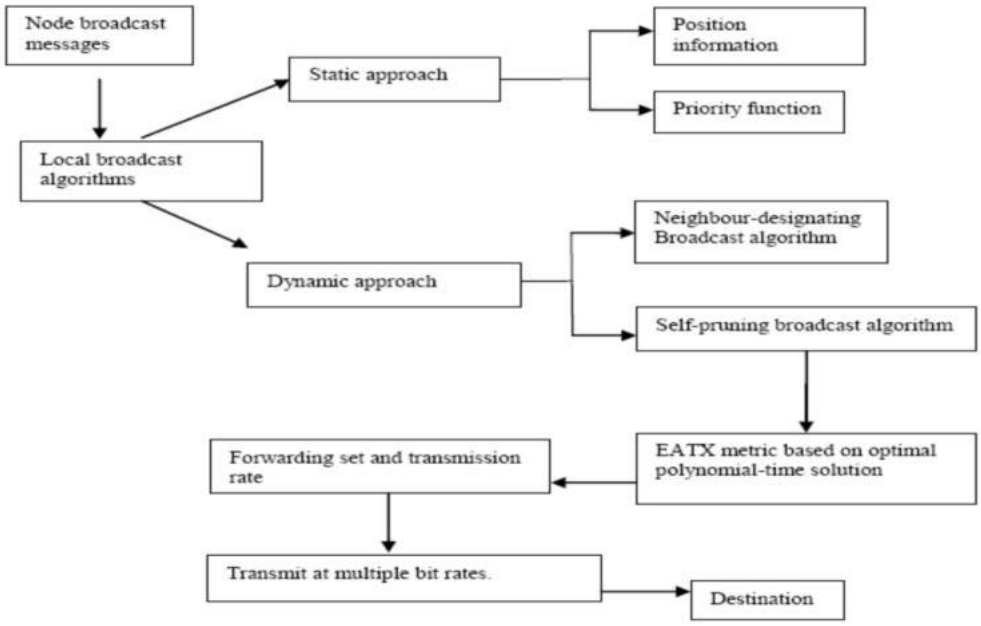
The hybrid (neighbour-designating and self-pruning) broadcast algorithm and show that the algorithm can achieve both full delivery and constant approximation only using connectivity information rather than position information. When a node u receives a message m, it creates a list satisfies coverage condition if it is not created yet and updates the list. Then based on whether u was selected to forward or whether the coverage condition is satisfied, node may schedule a broadcast by placing a copy of m in its MAC layer queue.

There are at least two sources of delay in the MAC layer. First, a message may not be at the head of the queue so it has to wait for other packets to be transmitted. Second, in contention based channel access mechanisms such as CSMA / CA to avoid collision, a packet at the head of the queue has to wait for a random amount of time before getting transmitted. The packet is removed from the MAC layer queue if it is no longer required to be transmitted. Therefore the broadcast algorithm has access to two functions to manipulate the MAC layer queue. The first function is the scheduling function which is used to place a message in the MAC layer queue. We assume that the scheduling function handles duplicate packets. i.e, it does not place the packet in queue if a copy of it is already in queue. Second function is called to remove a packet from the queue algorithm that needs to have access to MAC layer queue. This requires a cross-layer design. In absence of any cross-layer design the broadcast algorithm can use a timer at network layer

Algorithm: The proposed hybrid algorithm executed:

1. Extract ids of the broadcasting node and the selected node from the received message m
2. if u has broadcast the message m before then
3. Discard the message
4. Return
5. end if
6. if u receives m for the first time then
7. Create and fill the list $List_u^{cov}(m)$
8. end if
9. Update the list $List_u^{cov}(m)$
10. Remove the information added to the message by the previous broadcasting node
11. if $List_u^{cov}(m)!=\Phi$; then
12. Select an id from $List_u^{cov}(m)$ and add it to the message
13. Schedule the message {(*only update the selected id if m is already in the queue*)}
14. else {(*$List_u^{cov}(m)= \Phi$; in this case*)}
15. if u was selected then
16. Schedule the message {(*only remove the id of the selected neighbor if m is already in the queue*)}
17. else
18. Remove the message form the queue if u has not been selected by any node before
19. end if
20. end if

Figure-1: Architecture Diagram



Sources: Authors Compilation.

CONCLUSIONS AND FUTURE WORK

The local broadcast algorithm reduces the total number of transmissions that are required to achieve full delivery. Based on static approach it cannot guarantee a small sized CDS if the position information is not available.

It was shown that having relative position information can greatly simplify the problem of reducing the total number of selected nodes using the static approach.

Using local broadcast algorithms based on the dynamic approach require connectivity information rather than position information to guarantee a constant approximation factor.

The performance of multirate over single-rate anypath routing is analysed in our proposed system. Multirate outperforms single-rate anypath routing while still maintaining full connectivity to obtain optimum results.

REFERNCES

[1] B., Clark; C., Colbourn, and D., Johnson, (1990). "Unit Disk Graphs", Discrete Math., Volume: 86, pp. 165-177, 1990.

[2] M., Garey, and D., Johnson, (1990). "Computers and Intractability: A Guide to the Theory of NP-Completeness", W.H. Freeman & Co.

[3] J., Wu, and F., Dai, (2003). "Broadcasting in Ad Hoc Networks Based on Self-Pruning", Proceeding IEEE INFOCOM, pp. 2240-2250.

[4] J., Wu, and W., Lou, (2003). "Forward-Node-Set-Based Broadcast in Clustered Mobile Ad Hoc Networks", Wireless Comm. and Mobile Computing, Volume: 3, Number: 2, pp. 155-173.

[5] J., Wu, and F., Dai, (2004). "A Generic Distributed Broadcast Scheme in Ad Hoc Wireless Networks", IEEE Trans. Computers, Volume: 53, Number: 10, pp. 1343-1354, October.

*****

ALUMINIUM MATRIX COMPOSITES FOR AUTOMOBILE APPLICATIONS – A REVIEW

A. Thirumoorthy[19] T. V. Arjunan[20] S. Prakasam[21]

ABSTRACT

Metal composite materials have found application in many areas of daily life for quite some time. Often it is not realized that the application makes use of composite materials. These materials are produced in situ from the conventional production and processing of metals. The prospects of metal matrix composite technology and its note worthy application in the automotive industry is very high. The worldwide MMC markets in 2010 accounted for 2500 metric tons valued at over $1000M. Important MMC applications in the ground transportation (automobile and railway), thermal management, aerospace, industrial, recreational and infrastructure industries have been enabled by functional properties that include high structural efficiency, excellent wear resistance, and attractive thermal and electrical characteristics. Most of the researcher found out many new MMC materials for automobile industry and commercialized in various countries; because of their physical and mechanical properties made them amenable for use in bridge structures, particularly to retrofit existing damaged or deteriorated structural elements or as replacement of decks deteriorated by corrosion. In this review the alternate material for automobile brake applications with special attention to aluminum composites has been done.

KEYWORDS

AMC, Metal Matrix Composite, Strength-to-Weight Ratio, Weight etc.

INTRODUCTION

Weight reduction is one of the main objectives in automobile industry. In the last few years, considerable development has occurred in the potential use of cast metal matrix composites for automotive applications. Composite materials are usually classified on the basis of the physical or chemical nature of the matrix phase, e.g., polymer matrix, metal-matrix and ceramic composites. In addition there are some reports to indicate the emergence of Inter metallic matrix and carbon-matrix composite. Considerable trials have been done on use of metal matrix composites like aluminum silicon carbide composites for connecting rods, brake rotors, drive shafts and several other components. However, cost still remains a major barrier in designing aluminum composite components for wider applications in automotive industries. Aluminum and its alloys are attractive for many applications in chemical, automobile and aerospace industries because of their excellent properties as heigh strength-to-weight ratio, high electrical and thermal conductivities and good formability.

However their hardness, wear resistance and mechanical properties are poor in comparison to steel resistance and continuous efforts are made in the research into new possibilities for making use of the advantage of the aluminum in application that were reserved up to now for harder and more wear-resistant materials. This review is concerned with metal matrix composites and more specifically on the aluminium matrix composites (AMCs). In AMCs one of the constituent is aluminium/aluminium alloy, which forms percolating network and is termed as matrix phase. The other constituent is embedded in this aluminium/aluminium alloy matrix and serves as reinforcement, which is usually non-metallic and commonly ceramic such as SiC and Al2O3. Properties of AMCs can be tailored by varying the nature of constituents and their volume fraction.

Extensive studies have been carried out to assess the abrasion and sliding wear characteristics of AMCs. However, limited attempts were made on the solid particle erosion of these materials. Different investigators have reported different observations. Some reports suggested that AMCs suffer from higher erosive wear compared to the base alloy at impingement angles greater than 20 deg. On the other hand, others have reported that AMCs exhibit comparable erosive wear to that of the base alloy irrespective of the angle of impingement.

COMPOSITE MATERIALS

Composite materials signify that two or more constituents are combined on microscopic scales to synthesize a useful material. A variety of materials can be combined on a microscopic scale. The advantage of the composite materials is that their individual constituents retain their characteristic unlike alloys. As a result, various combinations of useful properties, usually not attainable by alloys, can be obtained through composite materials by suitable tailoring the matrix and reinforcement. In this case, other properties such as strength, hardness of the composites is less than that of the matrix alloy. However, they have been found to be within acceptable limits as confirmed through some experiments. The reinforcement of hard ceramic particles like silicon carbide, alumina, silica, zircon etc. in aluminium alloys has been found to improve the wear resistance as well as high temperature strength properties.

TYPES OF COMPOSITE MATERIALS

The most common composites can be divided into three main groups:

1. Polymer Matrix Composites (PMC's): Polymer matrix composites are also known as FRP - Fiber Reinforced Polymers (or Plastics). These materials use a polymer-based resin as the matrix, and a variety of fibers such as glass, carbon and aramid as the reinforcement.

2. Metal Matrix Composites (MMC's): Metal matrix composites are increasingly found in the aerospace and automotive industry. These materials use a metal such

[19]Research Scholar, Hi-Tech Engineering College, Tamil Nadu, India, a_tps@rediffmail.com
[20]Head and Professor, Coimbatore Institute of Engineering & Technology College, Tamil Nadu, India, arjun_nivi@yahoo.com
[21] Head and Professor, Shree Venkateshwara Hi-Tech Engineering College, Tamil Nadu, India, s_prakasam@rediffmail.com

as aluminum as the matrix, and reinforcement with fibers, particulates or whiskers such as silicon carbide.

3. Ceramic Matrix Composites (CMC's): Ceramic matrix composites are used in very high temperature environments. These materials use a ceramic as the matrix and reinforce it with short fibers, or whiskers such as those made from silicon carbide and boron nitride. Reductions in material density or increases in stiffness, yield strength, ultimate tensile strength can be directly translated to reductions in structural weight. This led the aerospace industry to develop and examine new materials with combinations of low density, improved stiffness and high strength as attractive alternatives to existing high-strength aluminum alloys and titanium alloys. These high-strength metal matrix composites combine the high strength and hardness of reinforcing phase with ductility and toughness of light metals. Moreover, the need for improved design procedures has resulted from an attempt to achieve significant improvement in structural efficiency, reliability and overall performance through reductions in either absolute weight or increases in strength-to-weight ratio. Recent research results have made it possible to envision combining these effects through the development of reinforced lightweight alloys.

The metal-matrix composites offer a spectrum of advantages that are important for their selection and use as structural materials. A few such advantages include the combination of high strength, high elastic modulus, high toughness and impact resistance, low sensitivity to changes in temperature or thermal shock, high surface durability, low sensitivity to surface flaws, high electrical and thermal conductivity, minimum exposure to the potential problem of moisture absorption resulting in environmental degradation, and improved fabric ability with conventional metal working equipment.

ALUMINUM MATRIX COMPOSITES (AMCS)

Aluminium alloys are used in advanced applications because their combination of high strength, low density, durability, machinability, availability and cost is very attractive compared to competing materials. However, the scope of these properties can be extended by using aluminium matrix composite materials. The aluminium matrix composites may offer specific advantages compared to unreinforced Al alloys, to polymer matrix composites and to ceramic matrix composites. Composite materials offer a unique opportunity to tailor the properties of aluminium. This could include increased strength, decreased weight, higher service temperature, improved wear resistance, higher elastic modulus, controlled coefficient of thermal expansion, improved fatigue properties. The fatigue properties of aluminium composites are usually better than the unreinforced equivalent alloys.

It is well recognized that when a soft metal like aluminum slides on hard steel without any external fluid or solid lubrication, the former is expected to flow and adhere to the latter, creating an interface of low shear strength. The transfer of aluminum on steel ball during a typical ball-on-disk friction test suppokrts this hypothesis. Transfer of aluminum will continue with sliding, and wear debris may form as a result of ploughing of the soft aluminum surface by the asperities of the hard steel, or flaking off of patches from the transfer film. Friction coefficient between aluminum and steel couples is high, 0.5–0.6. The development of aluminum MMCs dispersed with solid lubricants is primarily directed towards overcoming the principal drawbacks of aluminum as a tribological material. Rohatgi and coworkers first introduced graphite as a solid lubricant in aluminum matrices by casting routes, involving mixing the molten alloy with graphite particles to make a uniform suspension and followed by casting. The problem of graphite rejection by liquid aluminum is always faced here, caused by density differences (Al: 2.7 g/cc, graphite: 2.3 g/cc) and poor wettability between the two. These problems have been overcome to a large extent at the laboratory scale by the use of metal coatings (e.g., Ni and Cu) on the particle of graphite, and by the addition of reactive elements (e.g., Mg and Ti) to the melt. This work was subsequently extended to other solid lubricant dispersions but Al–Graphite by far has the most potential for commercial applications. There have been a number of publications in the literature on the sliding wear and friction of Al alloy–graphite composites. Unfortunately, different investigators have used different experimental parameters for hardness and roughness of the counter face, sliding speed, load and the test environment, making it difficult to quantify the effect of graphite content. In addition, comparing empirical wear data to theoretical generalizations of wear behavior is often difficult because of the widely different test conditions employed by different investigators to characterize the tribological properties of composites. In spite of a lack of universal testing procedure, useful generalizations concerning wear behavior of different materials, including composites, have been applied by constructing wear mechanism maps. Wear maps serve as predictive tools to draw meaningful conclusions relative to wear behavior under different test conditions. Specifically, some of the variations between different studies can be overcome by utilizing normalized test parameters such as non-dimensional wear rate, load, and sliding velocity.

Prasad and Rohatgi have introduced normalized wear rate (i.e., composite to the base matrix alloy) to analyze the data from different studies. It is a compilation of the normalized wear data on Al alloy–graphite composite showing the reduction in wear volume due to graphite particle dispersion. Compared with wear data, studies on friction behavior of Al–graphite composites are less numerous. Al alloy–graphite composites typically exhibit much lower friction coefficient compared with that of the matrix alloy. The coefficient of friction decreases considerably, up to about 3% by weight of graphite, and thereafter remains constant at about 0.2.

FABRICATION OF THE AMC'S

There are many processes viable to fabricate AMCs; they can be classified in: solid state, liquid-state and deposition processes.

In solid-state processes, the most spread method is powder metallurgy PM; it is usually used for high melting point matrices and avoids segregation effects and brittle reaction product formation prone to occur in liquid state processes. This method permits to obtain discontinuously particle reinforced AMCs with the highest mechanical properties. These AMCs are used for military applications but remain limited for large scale productions.

In liquid-state processes, one can distinguish the infiltration processes where the reinforcements form a preform which is infiltrated by the alloy melt: (1) With pressure applied by a

piston or by an inert gas (gas pressure infiltration GPI), and (2) without pressure. In the last case, one can distinguish: (a) the reactive infiltration processes using the wetting between reinforcement and melt obtained by reactive atmosphere, elevated temperature, alloy modification or reinforcement coating (reactive infiltration), and (b) the dispersion processes, such as stir-casting, where the reinforcements are particles stirred into the liquid alloy. Process parameters and alloys are to be adjusted to avoid reaction with particles. In deposition processes, droplets of molten metal are sprayed together with the reinforcing phase and collected on a substrate where the metal solidification is completed.

This technique has the main advantage that the matrix microstructure exhibits very fine grain sizes and low segregation, but has several drawbacks: the technique can only be used with discontinuous reinforcements, the costs are high, and the products are limited to the simple shapes that by obtained by extrusion, rolling or forging.

APPLICATIONS OF AMC

AMCs now have a proven track record as successful "high-tech" materials in a range of applications. AMC utilisation provides significant benefits including performance benefits (component lifetime, improved productivity), economic benefits (energy savings or lower maintenance cost) and environmental benefits (lower noise levels and fewer air-borne emissions). Engineering viability of AMCs in a number of applications have been well-documented. AMCs having different type of reinforcements (whiskers / particles / short fibres / continuous fibres) and produced both by solid state and liquid state processing have found their way too many practical applications.

Of all the AMCs, particle reinforced AMCs constitutes largest quantity of composites produced and utilised on volume and weight basis. PAMCs are produced by PM stir cast/melt infiltration/spraying/*in situ* processing techniques at industrial level. Particulates of SiC, Al_2O_3, TiC, TiB_2, B_4C have been used as reinforcements. PAMCs have been successfully used as components in automotive, aerospace, opto-mechanical assemblies and thermal management. PAMCs are in use as fan exit guide vane (FEGV) in the gas turbine engine, as ventral fins and fuel access cover doors in military aircraft. PAMCs are also used as rotating blade sleeves in helicopters. Flight control hydraulic manifolds made of 40 vol% SiCp reinforced aluminium composites have been successfully used.

The most notable large size and high volume use of PAMCs is in braking systems of trains and cars. Presently AMC brake discs are extensively used in European Railways and are in use in certain models of passenger cars in U.S.A. A major car manufacturing company in Europe is planning to introduce AMC brake discs by the year 2004. Potential automotive applications of PAMCs include valves, crankshafts, gear parts and suspension arms.

Particle reinforced AMCs are in use as recreational products including golf club shaft and head, skating shoe, base ball shafts, horseshoes and bicycle frames. AMCs containing high volume fraction ceramic particles are being used as microprocessor lids and integrated heat sinks in electronic packaging. They are also in use as carrier plates and microwave housing.

In the wake of greater health risks associated with the handling of ceramic whiskers, of late production of whisker-reinforced aluminium composites has been very limited.

However, with appropriate safety measures, SiC whisker-reinforced aluminium matrix composites have been produced and used as track shoes in advanced military tanks. Use of whisker-reinforced AMCs as track shoes help in reducing the weight of the tank. Short fibre-reinforced AMCs are being used in piston and cylinder liner applications.

Carbon fibre (continuous) reinforced Al matrix composites have been used as antenna wave guides for the Hubble Space Telescope. Here, composites provide high dimensional accuracy, high thermal and electrical conductivity with no out gassing oxidation resistance. 6061 Al-boron fibre (continuous) composites have been used as struts in main cargo bay of space shuttle. More recently 3MTm has developed alumina fibre (continuous) reinforced aluminium composites. Compared to high strength steel, CFAMCs developed by 3MTm offers equivalent strength at less than half the density and retains its strength to 300^0C and beyond. Composites possess four times the electrical conductivity of steel or half that of pure aluminium. These composites have been targeted for several functional applications including: (a) Core of an overhead electrical conductor; (b) automotive push rods; (c) flywheels for energy storage; (d) retainer rings for high-speed motors; and (e) brake calipers.

AMCs provide benefits in volume critical flywheel applications. Use of AMCs enables smaller flywheels compared to polymer composites. Thin walled retainer rings of AMCs provide excellent advantages in high-speed motors. AMC retainer rings can resist very high rotational speeds and still maintain their precise shape. Brake calipers made of CFAMCs offer the following advantages compared to cast iron brake calipers:

(1) Increased damping.
(2) Reduced unsprung weight.
(3) Increased fuel efficiency.
(4) Improved performance, handling and ride.
(5) Cross-platform utilisation.

AUTOMOTIVE BRAKES

Automotive brakes include drum and disk brakes for vehicle applications. Drum brakes predominantly use internally expanding shoes with brake linings that load the majority [typically 50 to 70%] of the drum rubbing surface. Most automotive disk brakes use shoes that load a much smaller portion [from 7 to 25%] of the disk rubbing surface. Disk brakes offer faster cooling, with their larger exposed surface areas and better cooling geometry, but are more vulnerable to either liquid or solid particulate contamination.

Rhee and others have related measurements on friction material wear with the brake cast iron temperature. These data indicated an essential constant wear rate for low temperature and a nearly exponential one at elevated temperature, because of cooling, contamination, and other basic design issues, front disk brakes and rear drum brakes are commonly used. The friction material is called a brake lining, whether it is used on drum. Drum brake linings are known as segments or strips, and heavy truck drum brake linings are called blocks. Particulate reinforced aluminum MMC s are promising candidate for automotive

applications since they offer high specific stiffness and strength, good wear resistance and suitable thermal properties; furthermore, they are readily available at reasonable prices and can be processed using conventional technologies.

A review of disc materials used by the automotive industry today will show that there are two basic material philosophies. The first, used for family sized vehicles, operates on the principal of small diameter, high strength, discs with sufficient inherent strength to resist any tendency towards the formation of thermal cracking, and distortion at high operating temperatures.

These discs whilst having good strength properties have relatively low thermal conductivity. The second principle, that of large weaker, low strength discs with high thermal conductivity, has been applied more commonly, to the larger high powered type of vehicle where space constraints are not so critical, and as a consequence, a large diameter thicker discs can be employed.

CONCLUSIONS

The properties of AMCs have been widely examined and would appear to offer several major advantages, namely:

- The friction coefficient of AMC is higher than that of cast Iron and better wear characteristics;
- The thermal conductivity of AMC can be two or three times higher than cast iron.

Clearly an impressive material, the performance of which depends on the nature

REFERENCES

[1] Kennedy. F.; E., Colin; F., Floquet, A., and Glovsky, R., (1984). "Improved Techniques for Finite Element Analysis of Sliding Surface Temperatures", Westbury House, pp. 138-150.

[2] Lin, J. Y., and Chen, H. T., (1992). "Radial Axis Symmetric Transient Heat Conduction in Composite Hollow Cylinders with Variable Thermal Conductivity", Volume: 10, pp. 2- 33.

[3] Brilla, J., (1997). "Laplace Transform and New Mathematical Theory of Visco Elasticity", Volume: 32, pp. 187- 195.

[4] Tsinopoulos, S. V.; Agnantiaris, J. P., and Polyzos, D., (1999). "An Advanced Boundary Element / Fast Fourier Transform Axis Symmetric Formulation for Acoustic Radiation and Wave Scattering Problems", Journal of ACOUST. SOC. AMER., Volume: 105, pp. 1517-1526.

[5] Wang, H. C., and Banerjee, P. K., (1990). "Generalized Axis Symmetri Elastodynamic Analysis by Boundary Element Method", Volume: 30, pp. 115-131.

[6] Floquet, A., and Dubourg, M. C., (1994). "Non Axis Symmetric Effects for Three Dimensional Analyses of a Brake", ASME Journal Tribology, Volume: 116, pp. 401-407.

[7] Burton, R. A., (1980). "Thermal Deformation in Frictionally Heated Contact", Wear, Volume: 59, pp. 1-20.

[8] Anderson, A. E., and Knapp, R. A., (1990). "Hot Spotting in Automotive Friction System", Wear, Volume: 135, pp 319-337.

[9] Comninou, M., and Dundurs, J., (1979). "On the Barber Boundary Conditions for Thermo Elastic Contact", ASME Journal, Volume: 46, pp. 849-853.

[10] Barber, J. R., (1978). "Contact Problems Involving a Cooled Punch", Journal of Elasticity, Volume: 8, pp. 409-423.

[11] Dow, T. A., and Burton, R. A., (1972). "Thermo Elastic Instability of Sliding Contact in Absence of Wear", Wear, Volume: 19, pp. 315-328.

[12] Lee, K., and Barber, J. R., (1993). "Frictionally-Excited Thermo Elastic Instability in Automotive Disk Brakes", ASME J. Tribology, Volume: 115, pp. 607-614.

[13] Lee, K., and Barber, J. R., Lee, (1994). "An Experimental Investigation of Frictionally excited Thermo Elastic Instability in Automotive Disk Brakes under a Drag Brake Application", ASME Journal of Tribology, Volume: 116, pp. 409-414.

[14] Lee, K., and Barber, J. R., Lee, (1995). "Effect of Intermittent Contact on the Stability of Thermo Elastic Contact", ASME Journal Tribology, Volume: 198, pp. 27-32.

[15] Zagrodzke, S.; Barber, J. R., and Hulbert, G. M., (1997). "Finite Element Analysis of Frictionally Excited Thermo elastic Instability", Journal Thermal Stresses, Volume; 20, pp. 185-201.

[16] Leroy, J. M.; Floquet, A., and Villechaise, B., (1989). 'Thermo-mechanical Behavior of Multilayered Media: Theory", ASME J. Tribology, Volume: 111, pp. 530-544.

[17] Gonska, H. W., and Kolbinger, H. J., (1993). "Temperature and Deformation Calculation of Passenger Car Brake Disks", Proceeding ABAQUS Users Conference, Aachen, Germany, pp. 21- 232.

[18] Akin, J. E., (1982). "Application and Implementation of Finite Element Methods", Academic Press, Orlando, FL, pp. 318-323.

[19] Zagrodzki, P., (1990). "Analysis of Thermo Mechanical Phenomena in Multi Disk Clutches and Brakes", Wear 140, pp. 291-308.

[20] Cook, R. D., (1981). Concept and Applications of Finite Element Analysis, Canada: Wiley.

[21] Zienkiewicz, O. C., (1977). The Finite Element Method, New York: McGraw-Hill.

[22] Beeker, A. A., (1992). The Boundary Element Method in Engineering, New York: McGraw-Hill.

*****

ENHANCING THE SAFETY OF AUTOMOBILE BY USING SHAPE MEMORY ALLOYS IN AIRBAGS

S. Ramesh Babu[22] R. Karthik[23]
A. Infant Augustine Vivek[24]

ABSTRACT

This paper deals with the method which enhances the safety of the automobile by using the shape memory alloys in the airbag system. The shape memory alloy is the special type of alloy having the property of regaining its original position after a deformation by applying some external heat source on it. The shape memory alloy has the tendency of regaining its own shape after deformation. The shape memory alloys are previously trained to respond to specified voltages to regain its position. The airbag is one of the most useful safety devices in automobiles. They inflate suddenly to safeguard the driver and passenger when the automobile hits on other vehicles or places.

The airbag is inflated by a geo storm unit controlled by the airbag controlling unit. As soon a the bumper of car hits the foreign body, the strain displacement sensor gives the impact force on the bumper and airbag controller unit orders airbag to inflate it. The average time of the inflation of airbag is 20 to 30 milliseconds; but it can inflate the airbag within 10 to 15 milliseconds by implementation of shape memory alloys. This will control the fatalities by 50% during future years and the safety of the drivers in the automobile will be ensured.

KEYWORDS

Shape-Memory Alloy, Pseudo Elasticity, Recrystallizes, Pyrotechnic etc.

SHAPE MEMORY ALLOYS

A shape-memory alloy is an alloy that "remembers" its original, cold-forged shape: returning the pre-deformed shape by heating. This material is a lightweight, solid-state alternative to conventional actuators such as hydraulic, pneumatic, and motor-based systems. Shape-memory alloys have applications in industries including medical and aerospace.

HISTORY

The first reported steps towards the discovery of the shape-memory effect were taken in the 1930s. According to Otsuka and Wayman, A. Ölander discovered the pseudo elastic behavior of the Au-Cd alloy in 1932. Greninger and Mooradian (1938) okbserved the formation and disappearance of a martensitic phase by decreasing and increasing the temperature of a Cu-Zn alloy.

The basic phenomenon of the memory effect governed by the thermoelastic behavior of the martensite phase was widely reported a decade later by Kurdjumov and Khandros (1949) and also by Chang and Read (1951).

The nickel-titanium alloys were first developed in 1962–1963 by the United States Naval Ordnance Laboratory and commercialized under the trade name Nitinol (an acronym for Nickel Titanium Naval Ordnance Laboratories). Their remarkable properties were discovered by accident. A sample that was bent out of shape many times was presented at a laboratory management meeting. One of the associate technical directors, Dr. David S. Muzzey, decided to see what would happen if the sample was subjected to heat and held his pipe lighter underneath it. To everyone's amazement the sample stretched back to its original shape.

There is another type of SMA, called a ferromagnetic shape-memory alloy (FSMA), that changes shape under strong magnetic fields. These materials are of particular interest as the magnetic response tends to be faster and more efficient than temperature-induced responses.

Metal alloys are not the only thermally-responsive materials; shape-memory polymers have also been developed, and became commercially available in the late 1990s.

CRYSTAL STRUCTURES

Many metals have several different crystal structures at the same composition, but most metals do not show this shape-memory effect. The special property that allows shape-memory alloys to revert to their original shape after heating is that their crystal transformation is fully reversible. In most crystal transformations, the atoms in the structure will travel through the metal by diffusion, changing the composition locally, even though the metal as a whole is made of the same atoms.

A reversible transformation does not involve this diffusion of atoms, instead the entire atoms shift at the same time to form a new structure, much in the way a parallelogram can be made out of a square by pushing on two opposing sides. At different temperatures, different structures are preferred and when the structure is cooled through the transition temperature, the martensitic structure forms from the austenitic phase.

TYPES OF SHAPE MEMORY ALLOYS

There are two types of shape memory alloys classified according to their remembrance capacity of their pre-deformed shape. They are:

- One way memory effect,
- Two way memory effect.

One-Way Memory Effect

When a shape-memory alloy is in its cold state (below A_s), the metal can be bent or stretched and will hold those shapes until heated above the transition temperature. Upon heating, the shape changes to its original. When the metal cools again it will remain in the hot shape, until deformed again.

[22] Associate Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, rathesh_28@yahoo.com
[23] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, karthik.mravi@gmail.com
[24] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, augustinvivek@gmail.com

With the one-way effect, cooling from high temperatures does not cause a macroscopic shape change. A deformation is necessary to create the low-temperature shape. On heating, transformation starts at A_s and is completed at A_f (typically 2 to 20 °C or hotter, depending on alloy or the loading conditions). A_s is determined by the alloy type and composition and can vary between −150 °C and 200 °C.

Two-Way Memory Effect

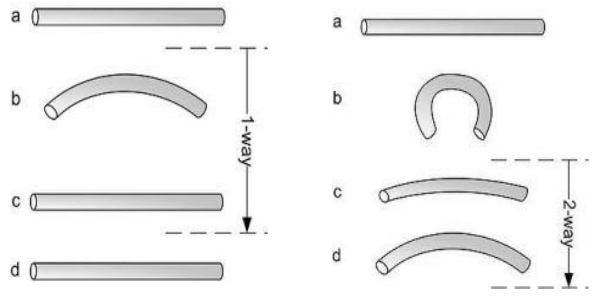
The two-way shape-memory effect is the effect that material remembers two different shapes: one at low temperatures, and one at the high-temperature shape. A material that shows a shape-memory effect during both heating and cooling is called two-way shape memory. This can also be obtained without the application of an external force (intrinsic two-way effect). The reason the material behaves so differently in these situations lies in training.

Training implies that a shape memory can "learn" to behave in certain way. Under normal circumstances, the shape - memory alloy "remembers" its high-temperature shape, but upon heating to recover the high-temperature shape, immediately "forgets" the low-temperature shape. However, it can be "trained" to "remember" to leave some reminders of the deformed low-temperature condition in the high-temperature phases. There are several ways of doing this. A shaped, trained object heated beyond a certain point will lose the two-way memory effect; this is known as "amnesia".

One-Way v/s Two-Way Shape Memory

Shape-memory alloys have different shape-memory effects. Two common effects are one-way and two-way shape memory. A schematic of the effects is shown below:

Figure-1



Sources: Authors Compilation.

The procedures are very similar: starting from martensite (a), adding a reversible deformation for the one-way effect or severe deformation with an irreversible amount for the two-way (b), heating the sample (c) and cooling it again (d).

PSEUDO-ELASTICITY

One of the commercial uses of shape-memory alloy exploits the pseudo-elastic properties of the metal during the high-temperature (austenitic) phase. The frames of reading glasses have been made of shape-memory alloy as they can undergo large deformations in their high-temperature state and then instantly revert back to their original shape when the stress is removed. This is the result of pseudo elasticity; the martensitic phase is generated by stressing the metal in the austenitic state and this martensite phase is capable of large strains. With the removal of the load, the martensite transforms back into the austenite phase and resumes its original shape.

This allows the metal to be bent, twisted and pulled, before reforming its shape when released. This means the frames of shape-memory alloy glasses are claimed to be "nearly indestructible" because it appears no amount of bending results in permanent plastic deformation.

The martensite temperature of shape-memory alloys is dependent on a number of factors including alloy chemistry. Shape-memory alloys with transformation temperatures in the range of 60–1450 K have been made.

MATERIALS

Alloys of metals having the memory effect at different temperatures and at different percentages of its solid solution contents:

- Ag-Cd 44/49 at.% Cd
- Au-Cd 46.5/50 at.% Cd
- Cu-Al-Ni 14/14.5 wt.% Al and 3/4.5 wt.% Ni
- Cu-Sn approx. 15 at.% Sn
- Cu-Zn 38.5/41.5 wt.% Zn
- Cu-Zn-X (X = Si, Al, Sn)
- Fe-Pt approx. 25 at.% Pt
- Mn-Cu 5/35 at.% Cu
- Fe-Mn-Si
- Pt alloys
- Co-Ni-Al
- Co-Ni-Ga
- Ni-Fe-Ga
- Ti-Pd in various concentrations
- Ni-Ti (~55% Ni)
- Ni-Ti-Nb
- Ni-Mn-Ga

MANUFACTURE

Shape-memory alloys are typically made by casting, using vacuum arc melting or induction melting. These are specialist techniques used to keep impurities in the alloy to a minimum and ensure the metals are well mixed. The ingot is then hot rolled into longer sections and then drawn to turn it into wire.

The way in which the alloys are "trained" depends on the properties wanted. The "training" dictates the shape that the alloy will remember when it is heated. This occurs by heating the alloy so that the dislocations re-order into stable positions, but not so hot that the material recrystallizes. They are heated to between 400 °C and 500 °C for 30 minutes. Typical variables for some alloys are 500 °C and for more than 5 minutes.

They are then shaped while hot and are cooled rapidly by quenching in water or by cooling with air.

PROPERTIES

The copper-based and NiTi-based shape-memory alloys are considered to be engineering materials. These compositions can be manufactured to almost any shape and size.

The yield strength of shape-memory alloys is lower than that of conventional steel, but some compositions have a higher yield strength than plastic or aluminum. The yield stress for Ni Ti can reach 500 MPa. The high cost of the metal itself and the processing requirements make it difficult and expensive to implement SMAs into a design. As a result, these materials are used in applications where the super elastic properties or the shape-memory effect can be exploited. The most common application is in actuation.

One of the advantages to using shape-memory alloys is the high level of recoverable plastic strain that can be induced. The maximum recoverable strain these materials can hold without permanent damage is up to 8% for some alloys. This compares with a maximum strain 0.5% for conventional steels.

APPLICATIONS

- Industrial,
- Aircraft,
- Piping,
- Automotive,
- Telecommunication,
- Robotics,
- Medicine,
- Optometry,
- Orthopedic surgery,
- Dentistry.

AIRBAG

An airbag is a vehicle safety device. It is an occupant restraint system consisting of a flexible envelope designed to inflate rapidly during an automobile collision. Its purpose is to cushion occupants during a crash and provide protection to their bodies when they strike interior objects such as the steering wheel or a window.

Modern vehicles may contain multiple airbags in various side and frontal locations of the passenger seating positions, and sensors may deploy one or more airbags in an impact zone at variable rates based on the type and severity of impact; the airbag is designed to only inflate in moderate to severe frontal crashes. Airbags are normally designed with the intention of supplementing the protection of an occupant who is correctly restrained with a seatbelt. Most designs are inflated through pyrotechnic means and can only be operated once.

The first commercial designs were introduced in passenger automobiles during the 1970s with limited success. Broad commercial adoption of airbags occurred in many markets during the late 1980s and early 1990s with a driver airbag, and a front passenger airbag as well on some cars; and many modern vehicles now include four or more units.

History: Invention

Reported in 1951, Walter Linderer designed an airbag. Linderer filed German patent #896,312 on October 6, 1951 which was issued on November 12, 1953, approximately three months after American John Hetrick was issued U.S. patent #2,649,311 earlier on August 18, 1953. Linderer's airbag was based on a compressed air system, either released by bumper contact or by the driver.

Later research during the 1960s proved that compressed air could not blow Linderer's airbag up fast enough for maximum safety, thus making it an impractical system.

John W. Hetrick, an industrial engineer and member of the United States Navy, designed the original safety cushion commonly referred to as an airbag. It was designed based on his experiences with compressed air from torpedoes during his service in the navy, as well as a need to provide protection for his family in their automobile during accidents. Hetrick worked with the major American automobile corporations at the time, but they chose not to invest in it.

In Japan, Yasuzaburou Kobori invented an airbag in 1963, on which technology current airbags are based, for which he was awarded patents in 14 countries. He died in 1975 without seeing widespread adoption of airbag systems.

In 1967, a breakthrough occurred in the development of airbags when Allen K. Breed invented a practical component for crash detection, an electromechanical sensor with a steel ball attached to a tube by a magnet that would inflate an airbag under a 30 milli-second window. Sodium azide instead of compressed air was also used for the first time during inflation. Breed Corporation then marketed this innovation first to Chrysler. A similar "Auto-Ceptor" crash-restraint, developed by Eaton, Yale & Towne Inc., for Ford was soon offered as an automatic safety system in the USA, while the Italian Eaton-Livia company offered a variant with localized air cushions.

As a Supplement to Seat Belts

Airbags for passenger cars were introduced in the United States in the mid-1970s, when seat belt usage rates in the country were quite low. Ford built an experimental fleet of cars with airbags in 1971, followed by General Motors in 1973 on Chevrolet vehicles. The early fleet of experimental GM vehicles equipped with airbags experienced seven fatalities, one of which was later suspected to have been caused by the airbag.

In 1974, GM made the ACRS or "Air Cushion Restraint System" available as a regular production option (RPO code AR3) in full-size Cadillac, Buick and Oldsmobile models. The GM cars from the 1970s equipped with ACRS have a driver side airbag, a driver side knee restraint (which consists of a padded lower dashboard) and a passenger side airbag. The passenger side airbag protects both front passengers and unlike most new ones, it integrates a knee cushion, a torso cushion and it also has dual stage deployment which varies depending on the force of the impact. The cars equipped with ACRS have lap belts for all seating positions but they do not have shoulder belts. These were already mandatory equipment in the United States on closed cars without airbags for the driver and outer front passenger seating positions.

The development of airbags coincided with an international interest in automobile safety legislation. Some safety experts advocated a performance-based occupant protection standard rather than a standard mandating a particular technical solution, which could rapidly become outdated and might not be a cost-effective approach. As countries successively mandated seat belt restraints, there was less emphasis placed on other designs for several decades. An airbag is not, and cannot be an alternative to seatbelts.

How Airbags Work?

The design is conceptually simple; a central "Airbag control unit" (ACU) monitors a number of related sensors within the vehicle, including accelerometers, impact sensors, side (door) pressure sensors, wheel speed sensors, gyroscopes, brake pressure sensors, and seat occupancy sensors. When the requisite 'threshold' has been reached or exceeded, the airbag control unit will trigger the ignition of a gas generator propellant to rapidly inflate a nylon fabric bag. As the vehicle occupant collides with and squeezes the bag, the gas escapes in a controlled manner through small vent holes. The airbag's volume and the size of the vents in the bag are tailored to each vehicle type, to spread out the deceleration of (and thus force experienced by) the occupant over time and over the occupant's body, compared to a seat belt alone.

The signals from the various sensors are fed into the Airbag control unit, which determines from them the angle of impact, the severity, or force of the crash, along with other variables. Depending on the result of these calculations, the ACU may also deploy various additional restraint devices, such as seat belt pre-tensioners, and/or airbags (including frontal bags for driver and front passenger, along with seat-mounted side bags, and "curtain" airbags which cover the side glass).

Each restraint device is typically activated with one or more pyrotechnic devices, commonly called an initiator or electric match. The electric match, which consists of an electrical conductor wrapped in a combustible material, activates with a current pulse between 1 to 3 amperes in less than 2 milliseconds. When the conductor becomes hot enough, it ignites the combustible material, which initiates the gas generator. In a seat belt pre-tensioner, this hot gas is used to drive a piston that pulls the slack out of the seat belt. In an airbag, the initiator is used to ignite solid propellant inside the airbag inflator.

The burning propellant generates inert gas which rapidly inflates the airbag in approximately 20 to 30 milliseconds. An airbag must inflate quickly in order to be fully inflated by the time the forward-traveling occupant reaches its outer surface. Typically, the decision to deploy an airbag in a frontal crash is made within 15 to 30 milliseconds after the onset of the crash, and both the driver and passenger airbags are fully inflated within approximately 60-80 milliseconds after the first moment of vehicle contact.

If an airbag deploys too late or too slowly, the risk of occupant injury from contact with the inflating airbag may increase. Since more distance typically exists between the passenger and the instrument panel, the passenger airbag is larger and requires more gas to fill it.

An airbag contains a mixture of sodium azide (NaN_3), KNO_3, and SiO_2. A typical driver-side airbag contains approximately 50-80 g of NaN_3, with the larger passenger-side airbag containing about 250 g. Within about 40 milliseconds of impact, all these components react in three separate reactions that produce nitrogen gas. The reactions, in order, are as follows:

- $2\ NaN_3 \rightarrow 2\ Na + 3\ N_2\ (g)$
- $10\ Na + 2\ KNO_3 \rightarrow K_2O + 5\ Na_2O + N_2\ (g)$
- $K_2O + Na_2O + 2\ SiO_2 \rightarrow K_2O_3Si + Na_2O_3Si$ (silicate glass)

The first reaction is the decomposition of NaN_3 under high temperature conditions using an electric impulse. This impulse generates to 300°C temperatures required for the decomposition of the NaN_3 which produces Na metal and N_2 gas. Since Na metal is highly reactive, the KNO_3 and SiO_2 react and remove it, in turn producing more N_2 gas.

The second reaction shows just that. The reason that KNO_3 is used rather than something like $NaNO_3$ is because it is less hygroscopic. It is very important that the materials used in this reaction are not hygroscopic because absorbed moisture can de-sensitize the system and cause the reaction to fail. The final reaction is used to eliminate the K_2O and Na_2O produced in the previous reactions because the first-period metal oxides are highly reactive. These products react with SiO_2 to produce a silicate glass which is a harmless and stable compound.

According to a patent, the particle size of the sodium azide, potassium nitrate, and silicon dioxide are important. The NaN_3 and KNO_3 must be between 10 and 20 microns, while the SiO_2 must be between 5 and 10 microns.

There has been a recent effort to find alternative compounds that can be used in airbags which have less toxic byproducts. In a journal article by Akiyoshi et. Al., it was found that for the reaction of the Sr complex nitrate, $(Sr(NH_2NHCONHNH_2)\cdot(NO_3)_2$ of carbohydrazide (SrCDH) with various oxidizing agents resulted in the evolution of N_2 and CO_2 gases.

Using $KBrO_3$ as the oxidizing agent resulted in the most vigorous reaction as well as the lowest initial temperature of reaction. The N_2 and CO_2 gases evolved made up 99% of all gases evolved. Nearly all starting materials won't decompose until reaching temperatures of 500°C or higher so this could be a viable option as an airbag gas generator.

In a patent containing another plausible alternative to NaN_3 driven airbags, the gas generating materials involved the use of guanidine nitrate, 5-amino tetrazole, bitetrazole dehydrate, nitroimidazole and basic copper nitrate. It was found that these non-azide reagents allowed for less toxic, lower combustion temperature reaction and more easily disposable airbag inflation system.

In vehicles equipped with a rollover sensing system, accelerometers and gyroscopes are used to sense the onset of a rollover event. If a rollover event is determined to be imminent, side-curtain airbags are deployed to help protect the occupant from contact with the side of the vehicle interior, and also to help prevent occupant ejection as the vehicle rolls over.

TRIGGERING CONDITIONS

Airbags are designed to deploy in frontal and near-frontal collisions more severe than a threshold defined by the regulations governing vehicle construction in whatever particular market the vehicle is intended for: U.S. regulations require deployment in crashes at least equivalent in deceleration to a 23 km/h (14 mph) barrier collision, or similarly, striking a parked car of similar size across the full front of each vehicle at about twice the speed. International regulations are performance based, rather than technology-based, so airbag deployment threshold is a function of overall vehicle design.

Unlike crash tests into barriers, real-world crashes typically occur at angles other than directly into the front of the vehicle, and the crash forces usually are not evenly distributed across the front of the vehicle. Consequently, the relative speed between a striking and struck vehicle required to deploy the airbag in a real-world crash can be much higher than an equivalent barrier crash. Because airbag sensors measure deceleration, vehicle speed is not a good indicator of whether an airbag should have deployed. Airbags can deploy due to the vehicle's undercarriage striking a low object protruding above the roadway due to the resulting deceleration.

The airbag sensor is a MEMS accelerometer, which is a small integrated circuit with integrated micro mechanical elements. The microscopic mechanical element moves in response to rapid deceleration, and this motion causes a change in capacitance, which is detected by the electronics on the chip that then sends a signal to fire the airbag. The most common MEMS accelerometer in use is the ADXL-50 by Analog Devices, but there are other MEMS manufacturers as well.

Initial attempts using mercury switches did not work well. Before MEMS, the primary system used to deploy airbags was called a "rolamite". A rolamite is a mechanical device, consisting of a roller suspended within a tensioned band. As a result of the particular geometry and material properties used, the roller is free to translate with little friction or hysteresis. This device was developed at Sandia National Laboratories. The rolamite and similar macro-mechanical devices were used in airbags until the mid-1990s when they were universally replaced with MEMS.

Nearly all airbags are designed to automatically deploy in the event of a vehicle fire when temperatures reach 150-200 °C (300-400 °F). This safety feature, often termed auto-ignition, helps to ensure that such temperatures do not cause an explosion of the entire airbag module.

Today, airbag triggering algorithms are becoming much more complex. They try to reduce unnecessary deployments and to adapt the deployment speed to the crash conditions. The algorithms are considered valuable intellectual property. Experimental algorithms may take into account such factors as the weight of the occupant, the seat location, seatbelt use, and even attempt to determine if a baby seat is present.

HOW IT IS IMPLEMENTED

The shape memory technology can be introduced in airbag system by weaving or synthesizing the shape memory alloy in the airbags. The suitable shape memory alloy is taken in thickness of the strand of the airbag wire. And it is made to be weaved or to be synthesized with the airbag material. The end of the shape memory alloy strand is connected to the cyclone creator of the airbag system.

To avoid the effect of heat on passenger is by adding elastic resin insulator on edges to expand along with the airbag to protect from heat. The resin insulating must be able to resist the heat and to be inert with the heat and the shape memory alloys. They should be able to have the usage to a longer extent.

ADVANTAGES

- It can be most effective to existing one.
- It reduces the inflating time of the airbag.
- It increases the safety of the airbag.
- Chances of fatalities can be decreased to an extent.
- It ensures the safety of peoples in vehicle

DISADVANTAGES

- The shape memory alloy can lose its property due to the high amount of heat on it.
- It can injure when it is not properly insulated with resin or elastic polymer.
- Complex in making the airbag material.

LIMITATIONS

- It may cost higher than existing one due to its complexity.
- It maybe varies in size with the existing model.

REFERENCES

[1]. K., Otsuka, and C., M. Wayman, Shape Memory Materials, Cambridge University Press.

[2]. Wu, S., and Wayman, C., (1987). "Martensitic transformations and the shape-memory effect in Ti50Ni10Au40 and Ti50Au50 alloys".

[3]. Filip, P., (1995). "Influence of Work Hardening and Heat Treatment on the Substructure and Deformation Behaviour of Tini Shape Memory Alloys", Scripta Metallurgica et. Materialia 32 (9): 1375.

[4]. Shape Memory Alloy Shape Training Tutorial.

[5]. Kauffman, George, and Isaac, Mayo, (1993). "Memory Metal", Chem Matters, October, pp. 4–7.

[6]. Oral history by William J. Buehler.

[7]. Huang, S., Leary, Martin; Attalla, Tamer; Probst, K., and Subic, A., (2012). "Optimisation of Ni–Ti shape Memory Alloy Response Time by Transient Heat Transfer Analysis", Materials & Design, 35: 655–663.

[8]. Leary, M.; Schiavone, F., and Subic, A., (2010). "Lagging for Control of Shape Memory Alloy Actuator Response Time", Materials & Design, 31 (4): 2124–2128.

[9]. The Lara Project – G1 and G2. Lararobot.de.

[10]. SMAterial.com – Phenomena, Crystallography, Model, Simulation and Applications of SMA – Has .gif animations demonstrating the effect.

*****

CLUSTERING WITH MULTIVIEWPOINT-BASED SIMILARITY MEASURE USING HIERARCHICAL CLUSTERING

Mayja Mary Varghese[25] Mookhambika Narayanan[26]

ABSTRACT

In data mining, clustering technique is an interesting and important technique. The main goal of the clustering is finding the similarity between the data points or similarity between the data within intrinsic data structure and grouping them the data into single groups (or) subgroups in clustering process.

The existing system is mainly used for finding the next frequent item set using k-Means method, where this algorithm can reduce the overlapping between the documents in itemset. The result of clustering process is based on the order for choosing the item sets in the k-Means approach; it doesn't follow a sequential order when selecting clusters. This problem will lead to gain less optimal solution for clustering method.

To resolve this problem, proposed system which is developing a novel hierarchal algorithm for document clustering produces superlative efficiency and performance. It is mainly focusing on making use of cluster overlapping phenomenon to design cluster merging criteria. Proposing a new way to compute the overlap rate in order to improve time efficiency and accuracy is mainly concentrated.

Based on the Hierarchical Clustering Method, the usage of Expectation-Maximization (EM) algorithm in the Gaussian Mixture Model is to count the parameters and make the two sub-clusters combined when their overlap is the largest is reported.

KEYWORDS

Hierarchical Clustering Method, Algorithum, Data Mining, Accuracy, Efficiency, Clustering, Performance, Expectation - Maximization etc.

INTRODUCTION

Cluster analysis or clustering is the task of assigning a set of objects into groups (called clusters) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative data mining, and a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics.

The concept of clustering has been around for a long time. It has several applications, particularly in the context of information retrieval and in organizing web resources. The main purpose of clustering is to locate information and in the present day context, to locate most relevant electronic resources. The research in clustering eventually led to automatic indexing to index as well as to retrieve electronic records. Clustering is a method in which we make cluster of objects that are somehow similar in characteristics. The ultimate aim of the clustering is to provide a grouping of similar records. Clustering is often confused with classification, but there is some difference between the two. In classification the objects are assigned to pre defined classes, whereas in clustering the classes are formed. The term "class" is in fact frequently used as synonym to the term "cluster". In database management, data clustering is a technique in which, the information that is logically similar is physically stored together. In order to increase the efficiency of search and the retrieval in database management, the number of disk accesses is to be minimized. In clustering, since the objects of similar properties are placed in one class of objects, a single access to the disk can retrieve the entire class.

When items in the database have been clustered, it is possible to retrieve all of the items in a cluster, even if the search statement did not identify them. When the user retrieves a strongly relevant item, the user can look at other items like it without issuing another search. When relevant items are used to create a new query (i.e., relevance feedback), the retrieved hits are similar to what might be produced by a clustering algorithm. However, term clustering and item clustering in a sense achieve the same objective even though they are the inverse of each other. The objective of both is to determine additional relevant items by a co-occurrence process.

For all of the terms within the same cluster, there will be significant overlap of the set of items they are found in. Item clustering is based upon the same terms being found in the other items in the cluster. Thus the set of items that caused a term clustering has a strong possibility of being in the same item cluster based upon the terms. For example, if a term cluster has 10 terms in it (assuming they are closely related), then there will be a set of items where each item contains major subsets of the terms. From the item perspective, the set of items that has the commonality of terms has a strong possibility to be placed in the same item cluster.

Hierarchical clustering techniques proceed by either a series of successive merges or a series of successive divisions. For both methods, the number of clusters is needed to select a clustering from the hierarchy. However the difference between the levels of the hierarchy may be an indication of the correct number of clusters. Hierarchical clustering generates a hierarchical tree of clusters. This tree is also called a dendrogram (Berkhin, 2005). Hierarchical methods can be further classified into agglomerative methods and divisive methods.

Agglomerative method: In an agglomerative method, originally, each object forms a cluster. Then the two most similar clusters are merged iteratively until some termination criterion is satisfied. Figure-1 shows an example for Hierarchical Agglomerative method.

Agglomerative algorithms are usually classified according to inter-cluster similarity measure they use. The most popular of these are single-link, complete-link and group average. Common for all agglomerative methods is high computational complexity, often quadratic or worse.

[25] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mayjamv@gmail.com
[26] Assitant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

Single-link clustering algorithms based on this similarity measure join the two clusters containing the two closest documents that are not yet in the same cluster. Complete-link clustering algorithms using this similarity measure join the two clusters with the minimum most-distant pair of documents. In this way, clusters are kept small and compact since all documents within a cluster have a maximum distance to each other (Frakes and Baeza-Yates, 1992). Group Average Clustering algorithms using this similarity measure join the two clusters with the minimum average document distance.

1.1 *K-means Clustering Algorithm*:

The K-means algorithm is by far the most popular clustering tool used in scientific and industrial applications. The name comes from representing each of K clusters C j where j ∈ {1, 2 …, K}, by the mean of its points, the so-called centroid. A centroid almost never corresponds to an actual data point.

The sum of differences between a point and its centroid expressed through appropriate distance is used as the objective function. It is the sum of the squares of errors between the points and the corresponding centroids, is equal to the total intra-cluster variance. This measure is used to show how well the centroids represent the members of their cluster.

$$E(C) = \sum_{j=1}^{K} \sum_{x_i \in C_j} \left| x_i - C_j \right|^2$$

K-means usually starts with selecting as initial cluster centers K randomly selected documents, the seeds. It then moves the cluster seed centers around in space in order to minimize E(C). This is done iteratively by repeating two steps until a stopping criterion is met: reassigning documents to the cluster with the closest centroid; and recomputing each centroid based on the current members of its cluster. Several termination conditions have been proposed:

- A fixed number of iterations have been completed;
- Assignment of documents to clusters does not change between iterations;
- Centroids do not change between iterations.

EXISTING SYSTEM

In the existing system, clustering is one of the most interesting and important topics in data mining. The aim of clustering is to find intrinsic structures in data, and organize them into meaningful subgroups for further study and analysis. There have been many clustering algorithms published every year. They can be proposed for very distinct research fields, and developed using totally different techniques and approaches.

Nevertheless, according to a recent study, more than half a century after it was introduced; the simple algorithm k-means still remains as one of the top 10 data mining algorithms nowadays. It is the most frequently used partitional clustering algorithm in practice.

Another recent scientific discussion states that k-means is the favorite algorithm that practitioners in the related fields choose to use. Needless to mention, k-means has more than a few basic drawbacks, such as sensitiveness to initialization and to cluster size, and its performance can be worse than other state-of-the-art algorithms in many domains. In spite of that, its simplicity, understandability, and scalability are the reasons for its tremendous popularity.

An algorithm with adequate performance and usability in most of application scenarios could be preferable to one with better performance in some cases but limited usage due to high complexity. While offering reasonable results, k-means is fast and easy to combine with other methods in larger systems. A common approach to the clustering problem is to treat it as an optimization process. An optimal partition is found by optimizing a particular function of similarity (or distance) among data. Basically, there is an implicit assumption that the true intrinsic structure of data could be correctly described by the similarity formula defined and embedded in the clustering criterion function.

Hence, effectiveness of clustering algorithms under this approach depends on the appropriateness of the similarity measure to the data at hand. For instance, the original k-means has sum-of-squared-error objective function that uses Euclidean distance. In a very sparse and high-dimensional domain like text documents, spherical k-means, which uses cosine similarity (CS) instead of Euclidean distance as the measure, is deemed to be more suitable.

PROPOSED SYSTEM

The proposed is motivated by investigations from the above and similar research findings. It appears to us that the nature of similarity measure plays a very important role in the success or failure of a clustering method. Our first objective is to derive a novel method for measuring similarity between data objects in sparse and high-dimensional domain, particularly text documents.

From the proposed similarity measure, we then formulate new clustering criterion functions and introduce their respective clustering algorithms, which are fast and scalable like k-means, but are also capable of providing high-quality and consistent performance. We develop two criterion functions for document clustering and their optimization algorithms. Finally extensive experiments on real-world benchmark data sets are presented.

We enhance our work to **Hierarchical Clustering Algorithms** given a set of N items to be clustered, and an N*N distance (or, similarity) matrix, the basic process of hierarchical clustering is this,

1. Start by assigning each item to a cluster, so that if you have N items, you now have N clusters, each containing just one item. Let the distances (similarities) between the clusters the same as the distances (similarities) between the items they contain.

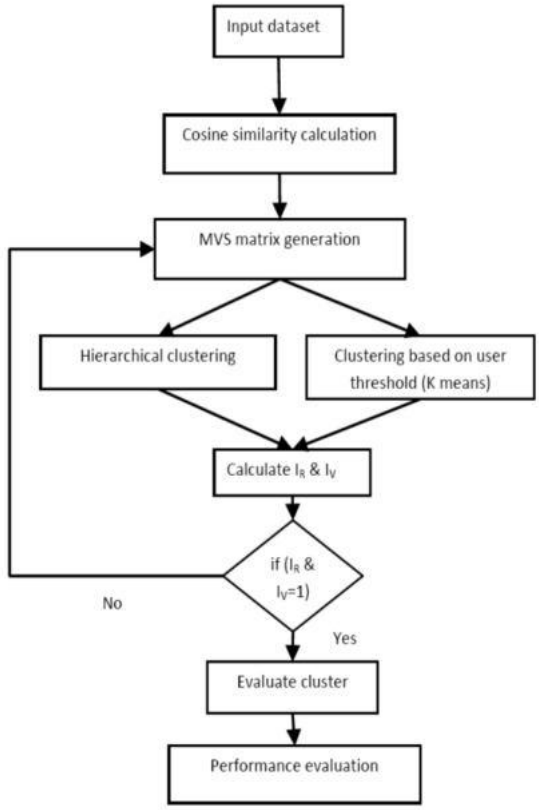
2. Find the closest (most similar) pair of clusters and merge them into a single cluster, so that now you have one cluster less.

3. Compute distances (similarities) between the new cluster and each of the old clusters.

4. Repeat steps 2 and 3 until all items are clustered into a single cluster of size N. (*)

Proposing a new way to compute the overlap rate in order to improve time efficiency and "the accuracy" is mainly concentrated. Experiments in both intra and inter of data and document clustering data show that this approach can improve the efficiency of clustering and save computing time. Given a data set satisfying the distribution of a mixture of Gaussians, the degree of overlap between components affects the number of clusters "perceived" by a human operator or detected by a clustering algorithm. In other words, there may be a significant difference between intuitively defined clusters and the true clusters corresponding to the components in the mixture.

Figure-1



Sources: Authors Compilation.

Term Frequency - Inverse Document Frequency

The TF-IDF is a text statistical-based technique which has been widely used in many search engines and information retrieval systems. Assume that there is a corporate of 1000 documents and the task is to compute the similarity between two given documents (or a document and a query). The following describes the steps of acquiring the similarity value:

Document Pre-Processing Steps:

- **Tokenization**: A document is treated as a string (or bag of words), and then partitioned into a list of tokens.
- **Removing Stop Words**: Stop words are frequently occurring, insignificant words. This step eliminates the stop words.
- **Stemming Word**: This step is the process of conflating tokens to their root form (connection -> connect).

Document Representation

- Generating N-distinct words from the corpora and call them as index terms (or the vocabulary). The document collection is then represented as N-dimensional vector in term space.

Measuring Similarity between 2 Documents

Capturing similarities of two documents using cosine similarity measurement. The cosine similarity is calculated by measuring the cosine of the angle between two document vectors. Using the code:

The main class is TFIDF Measure. This is the testing code:

```
void Test (string[] docs, int i, int j)
// docs is collection of parsed documents
{
StopWordHandler stopWord=new StopWordsHandler ();
TFIDFMeasure tf=new TFIDFMeasure (doc);
float simScore=tf.GetSimilarity (i, j); // similarity of two given documents at the // position i,j, respectively}
```

Simple Counting

This can begin by counting the number of times each of the words appear in each of the documents, Given this simple counting method, searches for "rose" can be sorted by its "term frequency" (TF) — the quotient of the number of times a word appears in each document (C), and the total number of words in the document (T) — TF = C / T. In the first case, rose has a TF value of 0.13. In the second case TF is 0.12, and in the third case it is 0.14. Thus, by this rudimentary analysis, Document 3 is most significant in terms of the word "rose", and Document 2 is the least. Document 3 has the highest percentage of content containing the word "rose".

TFIDF Analysis

By taking into account these two factors - term frequency (TF) and inverse document frequency (IDF) - it is possible to assign "weights" to search results and therefore ordering them statistically. Put another way, a search result's score ("ranking") is the product of TF and IDF: **TFIDF = TF * IDF** where:

- TF = C / T where C = number of times a given word appears in a document and T = total number of words in a document;
- IDF = D / DF where D = total number of documents in a corpus, and DF = total number of documents containing a given word;
- Given TFIDF, a search for "rose" still returns three documents ordered by Documents 3, 1, and 2.

A search for "newton" returns only two items ordered by Documents 2 (0.110) and 3 (0.061). In the later case, Document 2 is almost one and a half times more "relevant" than document 3. TFIDF scores can be summed to take into account Boolean unions (or) or intersections (and).

Cosine Similarity Measure

Cosine similarity is a measure of similarity between two vectors of *n* dimensions by finding the cosine of the angle between them, often used to compare documents in text mining. Given two vectors of attributes, *A* and *B*, the cosine similarity, θ, is represented using a dot product and magnitude as:

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}.$$

For text matching, the attribute vectors A and B are usually tf vectors of the documents. The cosine similarity can be seen as a method of normalizing document length during comparison.

Data pre-processing is an often neglected but important step in this project. In this module we are doing the preprocessing of two data sets of breast and wine. Analyzing data that has not been carefully screened for such problems can produce misleading results.

Thus, the representation and quality of data is first and foremost before running an analysis. If there is much irrelevant and redundant information present or noisy and unreliable data, then knowledge discovery during the training phase is more difficult. In order to improve the quality of data the preprocessing step done at first.

Data preparation and filtering steps can take considerable amount of processing time. Data pre-processing includes cleaning, normalization, transformation, feature extraction and selection, etc. The product of data pre-processing is the final training set. In this the parsing is the first step done when the document enters the process state. Parsing is defined as the separation or identification of Meta tags in a HTML document. Here, the raw HTML file is read and it is parsed through all the nodes in the tree structure.

The similarity between two documents is found by the cosine-similarity measure technique.

The weights in the cosine-similarity are found from the TF-IDF measure between the phrases (meta-tags) of the two documents.

This is done by computing the term weights involved:

TF = C / T,

IDF = D / DF.

D à quotient of the total number of documents,

DF à number of times each word is found in the entire corpus,

C à quotient of no of times a word appears in each document,

T à total number of words in the document,

TFIDF = TF * IDF

The cosine similarity in (3) can be expressed in the following form without changing its meaning:

$$Sim(d_i, d_j) = \cos(d_i - 0, d_j - 0) = (d_i - 0)^t (d_j - 0),$$

Clustering is a division of data into groups of similar objects. Representing the data by fewer clusters necessarily loses certain fine details, but achieves simplification. The similar documents are grouped together in a cluster, if their cosine similarity measure is less than a specified threshold.

It is used in the traditional k-means algorithm. The objective of k-means is to minimize the euclidean distance between objects of a cluster and that cluster's centroid:

$$Sim(d_i, d_j) = \cos(d_i, d_j) = d_i^t d_j.$$

Cosine measure is used in a variant of k-means called spherical k-means. While k-means aims to minimize euclidean distance, spherical k-means intends to maximize the cosine similarity between documents in a cluster and that cluster's centroid:

$$\max \sum_{r=1}^{k} \sum_{d_i \in S_r} \frac{d_i^t C_r}{\|C_r\|}.$$

Hierarchical clustering techniques proceed by either a series of successive merges or a series of successive divisions. For both methods, the number of clusters is needed to select a clustering from the hierarchy. However the difference between the levels of the hierarchy may be an indication of the correct number of clusters.

Hierarchical clustering generates a hierarchical tree of clusters. This algorithm is developing a novel hierarchal algorithm for document clustering. They used cluster overlapping phenomenon to design cluster merging criteria. The system computes the overlap rate in order to improve time efficiency and the veracity and the line passed through the two cluster's center instead of the ridge curve. Based on the hierarchical clustering method it used the Expectation-Maximization (EM) algorithm in the Gaussian mixture model to count the parameters and make the two sub-clusters combined when their overlap is the largest.

Having defined our similarity measure, we now formulate our clustering criterion functions. The first function, called IR, is the cluster size-weighted sum of average pairwise similarities of documents in the same cluster. First, let us express this sum in a general form by function F.

$$F = \sum_{r=1}^{k} n_r \left[\frac{1}{n_r^2} \sum_{d_i, d_j \in S_r} Sim(d_i, d_j) \right].$$

The final form of our criterion function I_R is:

$$I_R = \sum_{r=1}^{k} \frac{1}{n_r^{1-\alpha}} \left[\frac{n+n_r}{n-n_r} \|D_r\|^2 - \left(\frac{n+n_r}{n-n_r} - 1 \right) D_r^t D \right].$$

Again, we could eliminate n because it is a constant. Maximizing G is equivalent to maximizing Iv below;

$$I_V = \sum_{r=1}^{k} \left[\frac{n+\|D_r\|}{n-n_r} \|D_r\| - \left(\frac{n+\|D_r\|}{n-n_r} - 1 \right) \frac{D_r^t D}{\|D_r\|} \right].$$

Thus we have derived a novel method for measuring similarity between data objects in sparse and high-dimensional domain, particularly text documents. From the proposed similarity measure, we then formulate new clustering criterion functions

and introduce their respective clustering algorithms, which are fast and scalable like k-means, but are also capable of providing high-quality and consistent performance.

CONCLUSIONS

The experience in general data sets and a document set indicates that the new method can decrease the time cost, reduce the space complexity and improve the accuracy of clustering. In this paper, selecting different dimensional space and frequency levels leads to different accuracy rate in the clustering results.

We also developed an incremental insertion component for updating the comments-based hierarchy so that resources can be efficiently placed in the hierarchy as comments arise and without the need to re-generate the (potentially) expensive hierarchy.

REFERENCES

[1] X., Wu; V., Kumar; J., Ross Quinlan; J., Ghosh; Q., Yang; H., Motoda; G., J. McLachlan; A., Ng; B., Liu; P., S. Yu; Z., H. Zhou; M., Steinbach; D., J. Hand; D., Steinberg, (2007). "Top 10 algorithms in Data Mining", Knowl. Inf. Syst., Volume: 14, Number: 1, pp. 1–37.

[2] I., Guyon; U., Von Luxburg, and R. C. Williamson, (2009). "Clustering: Science or Art?, NIPS'09 Workshop on Clustering Theory.

[3] I., Dhillon, and D. Modha, (2001). "Concept Decompositions for Large Sparse Text Data Using Clustering", Mach. Learn., Volume: 42, Number: 1-2, pp. 143–175.

[4] S., Zhong, (2005). "Efficient Online Spherical K-means Clustering", in IEEE IJCNN, pp. 3180–3185.

[5] A., Banerjee; S., Merugu; I., Dhillon, and J. Ghosh, (2005). "Clustering with Bregman Divergences", Journal Mach. Learn. Res., October, Volume: 6, pp. 1705–1749.

[6] E., Pekalska; A., Harol; R., P. W. Duin; B., Spillmann; H., Bunke, (2006). "Non-Euclidean or Non-Metric Measures can be Informative", in Structural, Syntactic, and Statistical Pattern Recognition, ser. LNCS, Volume: 4109, pp. 871–880.

[7] M., Pelillo, (2009). "What is a cluster? Perspectives from game theory", in Proceeding of the NIPS Workshop on Clustering Theory.

[8] D., Lee, and J. Lee, (2010). "Dynamic Dissimilarity Measure for Support Based Clustering", IEEE Trans. on Knowl. and Data Eng., Volume: 22, Number: 6, pp. 900–905.

[9] A., Banerjee; I., Dhillon; J., Ghosh, and S. Sra, (2005). "Clustering on the Unit Hypersphere Using Von Mises-Fisher Distributions", J. Mach. Learn. Res., September, Volume; 6, pp. 1345–1382.

[10] W., Xu; X.m Liu, and Y. Gong, (2003). "Document Clustering Based on Non-negative Matrix Factorization", in SIGIR, pp. 267–273.

[11] H., Chin, and X. Deng, "Efficient Phrase-Based Document Similarity for Clustering".

*****

MINIMIZING DATA LOSS IN WIMAX SYSTEM USING BP ALGORITHM

V. Jaishankar[27] R. Sowmya[28] M. Nithya[29]

ABSTRACT

Many wireless communication systems such as IS-54, enhanced data rates for the GSM evolution , worldwide interoperability for microwave access (WiMAX) and Long Term Evolution (LTE) have adopted low-density parity-check tail-biting convolutional, and turbo codes as the Forward Error Correcting Codes (FEC) scheme for data and overhead channels. Therefore, many efficient algorithms have been proposed for decoding these codes. However, the different decoding approaches for these two families of codes usually lead to different hardware architectures. Since these codes work side by side in these new wireless systems, it is a good idea to introduce a universal decoder to handle these two families of codes. The present work exploits the Parity Check Matrix (H) representation of tail biting convolutional and turbo codes, thus enabling decoding via unified belief propagation (BP) algorithm. Indeed, the BP algorithm provides a highly effective general methodology for devising low-complexity iterative decoding algorithms for all convolutional code classes as well as turbocodes. While a small performance loss is observed when decoding turbocodes with BP instead of MAP, this is offset by the lower complexity of the BP algorithm and the inherent advantage of a unified decoding architecture.

KEYWORDS

WiMax, BP Algorithm, LTE, FEC, Parity Check Matrix, MAP etc.

I. INTRODUCTION

This document the convolutional codes were based on either algebraically calculating the error pattern or on trellis graphical representations such as in the AP and Viterbi algorithms. With the advent of turbo coding, a third decoding principle has appeared: iterative decoding. Iterative decoding was also introduced in Tanner's pioneering work , which is a general framework based on bipartite graphs for the description of LDPC codes and their decoding via the belief propagation (BP) algorithm. In many respects, convolutional codes are similar to block codes. For example, if we truncate the trellis by which a convolutional code is represented, a block code whose code words correspond to all trellis paths to the truncation depth is created.

However, this truncation causes a problem in error performance, since the last bits lack error protection. The conventional solution to this problem is to encode a fixed number of message blocks L followed by m additional all-zero blocks, where m is the constraint length of the convolutional code. This method provides uniform error protection for all information digits, but causes a rate reduction for the block code as compared to the convolutional code by the multiplicative factor L/(L+m). In the tail-biting convolutional code, zero-tail bits are not needed and replaced by payload bits resulting in no rate loss due to the tails. Therefore, the spectral efficiency of the channel code is improved.

Due to the advantages of tail-biting method over the zero-tail, it has been adopted as the FEC in addition to the turbo code for data and overhead channels in many wireless communications systems such as IS-54, EDGE, WiMAX and LTE. Both turbo and LDPC codes have been extensively studied for more than fifteen years. However, the formal relationship between these two classes of codes remained unclear until Mackay in claimed that turbo codes are LDPC codes. Also, Wiberg in marked another attempt to relate these two classes of codes together by developing a unified factor graph representation for these two families of codes. McEliece showed that their decoding algorithms fall into the same category as BP on the Bayesian belief network. Finally, Colavolpe was able to demonstrate the use of the BP algorithm to decode convolutional and turbo codes. The operation in is limited to specific classes of convolutional codes, such as convolutional self orthogonal codes (CSOCs). Also, the turbo codes therein are based on the serial structure while the parallel structure is more prevalent in practical applications.

II. FORWARD ERROR CORRECTION

FEC is accomplished by adding redundancy to the transmitted information using a predetermined algorithm. Each redundant bit is invariably a complex function of many original information bits. The original information may or may not appear in the encoded output; codes that include the unmodified input in the output are **systematic**, while those that do not are **non-systematic**.

An extremely simple example would be an analog to digital converter that samples three bits of signal strength data for every bit of transmitted data. If the three samples are mostly all zero, the transmitted bit was probably a zero, and if three samples are mostly all one, the transmitted bit was probably a one. The simplest example of error correction is for the receiver to assume the correct output is given by the most frequently occurring value in each group of three. This allows an error in any one of the three samples to be corrected by "democratic voting". This is a highly inefficient FEC, and in practice would not work very well, but it does illustrate the principle.

In practice, FEC codes typically examine the last several dozen, or even the last several hundred, previously received bits to determine how to decode the current small handful of bits (typically in groups of 2 to 8 bits).Such triple modular redundancy, the simplest form of forward error correction, is widely used.Forward Error Correction (FEC) is a type of error correction which improves on simple error detection schemes by enabling the receiver to correct errors once they are detected. This reduces the need for retransmissions. FEC works by adding check bits to the outgoing data stream. Adding more check bits reduces the amount of available bandwidth, but also enables the receiver to correct for more errors. Forward Error Correction is particularly well suited for satellite transmissions, where bandwidth is reasonable but latency is significant.

[27] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, jaishankarmecse@gmail.com
[28] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, Sowmyaraju.yes@gmail.com
[29] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, nithi.mn@gmail.com

2.1. Averaging Noise to Reduce Errors

FEC could be said to work by "averaging noise"; since each data bit affects many transmitted symbols, the corruption of some symbols by noise usually allows the original user data to be extracted from the other, uncorrupted received symbols that also depend on the same user data.

- Because of "risk-pooling" effect, digital communication systems that use FEC tend to work perfectly above a certain minimum signal-to-noise ratio and not at all below it.
- This all-or-nothing tendency becomes more pronounced as stronger codes are used that more closely approach the theoretical limit imposed by the Shannon limit.
- Interleaving FEC coded data can reduce the all or nothing properties of transmitted FEC codes. However, this method has limits. It is best used on narrowband data.

Most telecommunication systems used a fixed channel code designed to tolerate the expected worst-case bit error rate, and then fail to work at all if the bit error rate is ever worse.

However, some systems adapt to the given channel error conditions: hybrid automatic repeat-request uses a fixed FEQ method as long as the FEQ can handle the error rate, then switches to ARQ when the error rate gets too high; adaptive modulation and coding uses a variety of FEQ rates.

2.2. Types of FEC

The two main categories of FEC are: a) Block Coding, and b) Convolutional Coding.

- Block codes work on fixed-size blocks (packets) of bits or symbols of predetermined size.
- Convolutional codes work on bit or symbol streams of arbitrary length. A convolutional code can be turned into a block code, if desired. Convolutional codes are most often decoded with the Viterbi algorithm, though other algorithms are sometimes used.

There are many types of block codes, but the most notable is Reed-Solomon coding because of its widespread use on the Compact disc, the DVD, and in computer hard drives. Golay, BCH and Hamming codes are other examples of block codes. Hamming ECC is commonly used to correct NAND flash memory errors. This provides single-bit error correction and 2-bit error detection.

Hamming codes are only suitable for more reliable single level cell (SLC) NAND. Denser multi level cell (MLC) NAND requires stronger multi-bit correcting ECC such as BCH or Reed-Solomon·Nearly all block codes apply the algebraic properties of finite fields.

2.3. Concatenate Fec Codes to Reduce Errors

Block and convolutional codes are frequently combined in **concatenated** coding schemes in which the convolutional code does most of the work and the block code (usually Reed-Solomon) "mops up" any errors made by the convolutional decoder. This has been standard practice in satellite and deep space communications since Voyager 2 first used the technique in its 1986 encounter with Uranus.

III. TURBO CODES

The most recent (early 1990s) development in error correction is turbo coding, a scheme that combines two or more relatively simple convolutional codes and an interleaver to produce a block code that can perform to within a fraction of a decibel of the Shannon limit.

- One of the earliest commercial applications of turbo coding was the CDMA2000 1x (TIA IS-2000) digital cellular technology developed by Qualcomm and sold by Verizon Wireless, Sprint, and other carriers.
- The evolution of CDMA2000 1x specifically for Internet access, 1xEV-DO (TIA IS-856), also uses turbo coding. Like 1x, EV-DO was developed by Qualcomm and is sold by Verizon Wireless, Sprint, and other carriers (Verizon's marketing name for 1xEV-DO is Broadband Access, Sprint's consumer .

IV. INTERLEAVING

Interleaving in computer science is a way to arrange data in a non-contiguous way in order to increase performance. It is used in:

- Time Division Multiplexing (TDM) in telecommunication.
- Computer memory.
- Disk storage.

Interleaving is mainly used in data communication, multimedia file formats, radio transmission (for example in satellites) or by ADSL. The term multiplexing is sometimes used to refer to the interleaving of digital signal data.

4.1. Interleaving In Disk Storage

Low-level format utility performing interleaves speed tests on a 10-megabyte IBM PC XT hard drive.Historically, interleaving was used in ordering block storage on disk-based storage devices such as the floppy disk and the hard disk. The primary purpose of interleaving was to adjust the timing differences between when the computer was ready to transfer data, and when that data was actually arriving at the drive head to be read. Interleaving was very common prior to the 1990s, but faded from use as processing speeds increased. Modern disk storage is not interleaved. Interleaving was used to arrange the sectors in the most efficient manner possible, so that after reading a sector, time would be permitted for processing, and then the next sector in sequence is ready to be read just as the computer is ready to do so. Matching the sector interleave to the processing speed therefore accelerates the data transfer, but an incorrect interleave can make the system perform markedly slower.

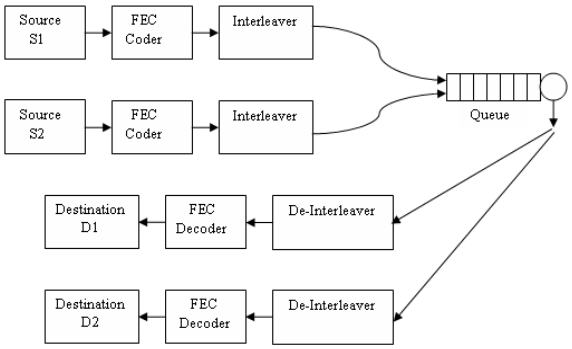
4.2. Interleaving In Data Transmission

Interleaving is used in digital data transmission technology to protect the transmission against burst errors. These errors overwrite a lot of bits in a row, so a typical error correction scheme that expects errors to be more uniformly distributed can be overwhelmed. Interleaving is used to help stop this from happening. Data is often transmitted with error control bits that enable the receiver to correct a certain number of errors that occur during transmission. If a burst error occurs, too many errors can be made in one code word, and that codeword cannot be correctly decoded. To reduce the effect of such burst errors,

the bits of a number of codewords are interleaved before being transmitted. This way, a burst error affects only a correctable number of bits in each codeword, and the decoder can decode the codewords correctly. This method is popular because it is a less complex and cheaper way to handle burst errors than directly increasing the power of the error correction scheme.

V. SYSTEM DESIGN

Figure-1



Sources: Authors Compilation.

5.1. FEC Ecoder

FEC is a system of error control for data transmission, where the sender adds redundant data to its messages. This allows the receiver to detect and correct errors (within some bounds) without the need to ask the sender for additional data. In this module we add redundant data to the given input data, known as FEC Encoding. The text available in the input text file is converted into binary. The binary conversion is done for each and every character in the input file. Then we add the redundant data for each bit of the binary. After adding we have a block of packets for each character. The User Interface design is also done in this module. We use the Swing package available in Java to design the User Interface. Swing is a widget toolkit for Java. It is part of Sun Microsystems' Java Foundation Classes (JFC) — an API for providing a graphical user interface (GUI) for Java programs.

5.2. Interleaver

Interleaving is a way of arranging data in a non-contiguous way in order to increase performance. It is used in data transmission to protest against burst errors. In this module we arrange the data (shuffling) to avoid burst errors which is useful to increase the performance of FEC Encoding. It gets the input as blocks of bits from the FEC Encoder. In this module we shuffle the bits inside a single block in order to convert burst errors into random errors. This shuffling process is done for each and every block comes from the FEC Encoder. Then we create a Socket connection to transfer the blocks from Source to the Queue. This connection is created by using the Server Socket and Socket class Available in Java.

5.3. Implementation of the Queue

We receive the data from the Source system. This data is the blocks after FEC Encoding and Interleaving processes are done. These blocks come from the Source system through Server Socket and Socket. Server socket and Socket are classes available inside Java. These two classes are used to create a connection between two systems inside a network for data transmission. After we receive the packets from Source, we create packet loss. Packet loss is a process of deleting the packets randomly. After creating loss we send the remaining blocks to the Destination through the socket connection.

5.4. De-Interleaver

De-Interleaver receives the blocks of data from the Queue through the socket connection. These blocks are the remaining packets after the loss in the Queue. In this module we re arrange the data packets inside a block in the order in which it is before Interleaving. This process of Interleaving and De-Interleaving is done to convert burst errors into random errors. After De-Interleaving the blocks are arranged in the original order. Then the data blocks are sent to the FEC Decoder.

5.5. FEC Decoder

FEC decoder gets the input from the De-Interleaver. The received packets are processed to remove the original bits from it. Thus we recover the original bits of a character in this module. After retrieving original bits, we convert it to characters and write it inside a text file. We calculate the overall performance of FEC Coding in recovering the packet losses. After retrieving the original bits, we convert it to characters and write it inside a text file. This performance is calculated by using coding parameters like Coding rate, Interleaving depth, Block length and several other parameters. First we calculate the amount of packet loss and with it we use various formulas to calculate overall performance of Forward Error Correction in recovering the network packet losses.

VI. WIMAX AND LTE CODING STRUCTURE

To address the low and high rate requirements of LTE, the 3rd Generation Partnership Project (3GPP) working group under took a rigorous study of advanced channel coding candidates such as tail-biting convolutional and turbo codes for low and high data rates, respectively. We investigate here the application of the BP decoder for the proposed turbo code in LTE systems. Meanwhile, a rate ½, memory-6 tail-biting convolutional code has been adopted in the WiMAX (802.16*e*)system, because of its best minimum distance and the smallest number of minimum weight codewords for larger than 32-bitpayloads which is used for both frame control header (FCH) and data channels.

VII. DESIGN OBJECTIVES

The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves system's relationship to help user decision-making:

1) Designing computer output should proceed in an organized, well thought out manner; the right output

must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2) Select methods for presenting information.
3) Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives:

- Convey information about past activities, current status or projections of the future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

VIII. CONCLUSIONS

In this paper, the feasibility of decoding arbitrary tailbiting convolutional and turbo codes using BP algorithm was demonstrated. Using this algorithm to decode the tailbiting convolutional code in WiMAX systems speeds up the error correction convergence and reduces decoding computational complexity with respect to the ML-Viterbi-based algorithm. In addition, the BP algorithm performs a non-trellis based forward-only algorithm and has only an initial decoding delay, thus avoiding intermediate decoding delays that usually accompany the traditional MAP and SOVA components in LTE turbo decoders. However, with respect to the traditional decoders for turbo codes, the BP algorithm is about 1.7 dB worse at a BER value of 10−2. This is because the nonzero element distribution in the parity-check matrix is not random enough. Also, there are number of short cycles in corresponding Tanner graphs. Finally, as an extended work, we propose the BP decoder for these codes in a combined architecture which is advantageous over a solution based on two separate decoders due to efficient reuse of computational hardware and memory resources for both decoders

IX. REFERENCES

[1] C., Berrou; A., Glavieux, and P., Thitimajshima, (1993). "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes", IEEE International Conference on Commun., Volume-2, Geneva, Switzerland, pp. 1064-1070.

[2] R., M. Tanner, (1981). "A Recursive Approach to Low Complexity Codes", IEEE Trans. Inform. Theory, Volume: 27, No. 5, pp. 533-547.

[3] G., D. Forney, Jr., (2001). "Codes on Graphs: Normal Realizations", IEEE Trans. Inform. Theory, Volume: 47, Number: 2, pp. 520–548.

[4] H., H. Ma, and J., K. Wolf, (1986). "On Tail Biting Convolutional Codes", IEEE Trans. On Commun., Volume: 34, Number: 2, pp. 104-111.

[5] 3GPP TS 45.003, "3rd Generation Partnership Project; Technical Specification Group GSM/EDGE Radio Access Network; Channel Coding (Release 7)", February, 2007.

[6] IEEE Std 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems," October, 2004.

[7] IEEE Std P802.16e/D10, "IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems," August, 2005.

[8] D., J. C. MacKay, (1999) "Good error-correcting codes based on very sparse matrices", IEEE Trans. Inform. Theory, Volume: 45, Number: 2, pp. 399-431.

[9] N., Wiberg, (1996). "Codes and Decoding on General Graphs", Linkoping Studies in Science and Technology, Dissertation No. 440, Linkoping University, Linkoping, Sweden.

[10] R., J. McEliece; D., MacKay, and J. Fu Cheng, (1998). "Turbo Decoding as an Instance of Pearl's "Belief Propagation" Algorithm", IEEE Trans. On Commun, Volume: 16, Number: 2, pp. 140-152.

[11] G., Colavolpe, (2004), "Design and performance of turbo Gallager codes," IEEE Trans. On Commun., Volume: 52, Number: 11, pp. 1901-1908.

[12] T., T. Chen, and S-He Tsai, (2008). "Reduced-Complexity Wrap-Around Viterbi Algorithms for Decoding Tail-Biting Convolutional Codes", 14th European Wireless Conference, Junuary.

[13] M., G. Luby; M., Mitzenmacher; M., A. Shokrollahi, and D., A. Spielman, (2001). "Improved LDPC Codes Using Irregular Graphs", IEEE Trans. Inform. Theory, Volume: 47, Number; 2, pp. 585-598.

[14] Giulietti A., (2004). "Turbo Codes: Desirable and Designable", IKluwer Academic Publishers, ISBN: 1-4020-7660-6.

[15] P., Elias, (1955). "Coding for Noisy Channels", IRE Conv. Rec., Volume: 3, Pt. 4, pp. 37–46.

[16] A., J. Viterbi, (1971). "Convolutional Codes and Their Performance in Communication Systems", IEEE Trans. On Commun., Volume: 19, Number: 15, pp. 751-772.

[17] C., Poulliat; D., Declercq, and T., Lestable, (2008). "Efficient Decoding of Turbo Codes with Nonbinary Belief Propagation" EURASIP Journal on Wireless Communications and Networking, Number: 473613.

[18] A., Refaey; S., Roy, and P., Fortier, (2009). "On the Application of BP Decoding to Convolutiona and Turbo Codes", Asilomar Conference on Signals, Systems, and Computers, November.

[19] P., Stahl; J., B. Anderson, and R., Johannesson, (1999). "Optimal and Near-Optimal Encoders for Short and Moderate-Length Tail-Biting Trellises", IEEE Trans. Inform. Theory, Volume: 45, Number: 7, pp. 2562-2571.

[20] W., Yu; M., Ardakani; B., Smith, and F., FKschischang, "Complexityoptimized low-density parity-check codes for Gallager decoding algorithm B", IEEE International Symposium on Information Theory (ISIT)s.

*****

CATALYST BASED EXTRACTION OF PYROLYSIS OIL FROM ACACIA ARABICA SEED

K. Venkatesan [30] M. Senthilkumar[31]

ABSTRACT

The materials used by human beings for their life and the disposal of used materials or residues from waste cause pollution to the environment. Pollution or contamination of these wastes which is making problems for living and nonliving organisms in the ecosystem. The types of waste are solid, liquid, gaseous, and radioactive. Waste management is carrying out to dispose the materials. Waste management involves the process of collection, transport, processing, recycling and disposal. The disposal of solid wastes particularly by preparing land fill, incineration, pumping into deep wells and burning does not solve the problem. Changing the solid waste into some useful form like alternative fuel will help to address the energy crises. This project work focuses on extracting oil from various biomasses at yield temperature. Pyrolysis is a thermochemical decomposition of an organic material at elevated temperature in absence of oxygen. Pyrolysis process is used to extract the bio oil from various biomasses. Catalysts like KOH, $CaCO_3$, Nacl is selected according to its oxidation characteristics and is added with acacia arabica. Then pyrolysis process is carried out at a temperature of 620 -700˚C. As a result of adding the catalyst, there is an increase in the extraction of oil. The increase in extraction of oil product quantity is compensated mainly with decrease in gas product. The objective of work is to determine the effect of reaction, temperature and time on optimization of pyrolysis to produce renewable bio oil. Of three catalysts, Nacl alone has the greatest power of the extraction of oil from biomass. The optimum oil product yield of 20% is achieved at 620˚C by using the Nacl.

KEYWORDS

Catalysts, Biomass, Waste, Pyrolysis. Acacia Arabica, Sodium Chloride (Nacl), Potassium Hydroxide (KOH), Calcium Carbonates ($CaCO_3$) etc.

1. INTRODUCTION

Renewable energy is of growing importance in satisfying environmental concerns over fossil fuel usage. Wood and other forms of biomass including energy crops and agricultural and forestry wastes are some of main renewable energy resources available. These can provide the only source of renewable liquid, gaseous and solid fuels. Bio mass is considered the renewable energy source with the highest potential to contribute to the energy needs of modern society for both the developed and developing economics world-wide. Bio energy is now accepted as having the potential to provide the major part of the projected renewable energy provisions of the future. Pyrolysis is one of the three main thermal routes, with gasification and combustion, to providing a useful and valuable bio-fuel. It is one of the most recent renewable energy processes to have been introduced and offers the advantages of a liquid product bio-oil that can be readily stored and transported, and used as a fuel, an energy carrier and a source of chemicals. Pyrolysis has now achieved commercial success for production of some chemicals, liquid fuel and electricity. Bio-oils have been successfully tested in engines, turbines and boilers, and have been up graded to high quality hydro carbon fuels although at a presently unacceptable energetic and financial cost. This review concentrates on the technology of pyrolysis and application for the liquid product.

2. PRODUCT OF PYROLYSIS

2.1 Syngas

Syngas, or synthesis gas, refers to a gas mixture that contains different amounts of carbon monoxide (CO) and hydrogen, produced from the gasification of biomass or fossil fuels. Syngas can be burned directly to produce heat or power.

- It can be converted through a chemical process to synthetic natural gas (SNG).
- Syngas can be converted into a wide range of liquid bio- fuels including synthetic diesel, methanol, DME, and others.
- Processes that convert biomass into liquid fuels, using syngas as an intermediate step are often called biomass – to – liquids or BTL.

2.2 Bio-Oil

Bio oil is a dark brown liquid produced from a pyrolysis process. It is composed of a complex mixture of oxygenated hydrocarbons with an appreciable proportion of water. The oil may contain some solid char. The easiest way to the fuel is to use it in combustion process. Bio oil would have a higher value if it could be used as engine fuel for cars and trucks.Alternatively the oil can be upgraded to either special engine fuel or converted into a syngas through a gasification process and thereafter to bio- fuel.

2.3 Char

The de-volatization of biomass during pyrolytic reactions yields a solid residue (char). Increasing the heat treatment temperature reduces char yield and also increase the aromatization of char as measures by the aromatic carbon content of the zcids. Char is most commonly used as a fertilizeror soil amendment. Char may also have the potential to sequester large amounts of carbon in the soil.

3. PROPERTIES OF BIOMASS PYROLYSIS OILS

Pyrolysis liquid is referred to by many names including pyrolysis oil, bio- oil, bio crude oil, bio-fuel-oil, wood liquids, wood oil, liquid smoke, wood distillates, pyroligneous tar, pyroligneous acid, and liquid wood.

The crude pyrolysis liquid is usually dark brown and free flowing with a distinctive smoky smell. Chemically, it approximates to bio mass in elemental composition and is

[30]Research Scholar, S.K.P. Engineering College, Tamil Nadu, India, adervenkatesan1970@yahoo.com
[31]Assistant Professor, Madras Institute of Technology, Tamil Nadu, India, mskiitm@yahoo.com

composed of a very complex mixture of oxygenated hydro carbons with an appreciable proportion of water from both the original moisture and reaction product. Solid char may also be present. The physical properties of bio-oils are described in several publications. These properties result from the chemical composition of the oils, which is significantly different from that of petroleum derived oils. Bio-oils are multi-component mixtures comprised of different size molecules derived primarily from depolymerization and fragmentation reactions of three key bio mass building blocks: cellulose, hemi cellulose, and lignin. Therefore, elemental composition of bio-oil resembles that of bio mass rather than that of petroleum oil.

3.1 Catalysts

Catalyst is the material which is used to induce reaction and fast the speed of the reaction.

3.2 Types of Catalyst

In these projects different types of catalyst are used:

1. Sodium chloride (Nacl)
2. Potassium hydroxide (KOH)
3. Calcium carbonate ($CaCO_3$)

4. PROPERTIES OF CATALYST

Sodium Chloride (Nacl)

- Removes water from oil.
- Increases dissociation of organic matter by its polar structure.

Potassium Hydroxide (KOH)

- Removes carbon dioxide.
- It can also control emission of CO.

Calcium Carbonate ($CaCo_3$)

- To maintain moderate temperature.
- It can easily helps to removes moisture and volatile matters

5. EXPERIMENTAL PROCEDURE

The basic steps followed in the experimentation are:

- The samples such as acacia arabica are prepared for pyrolysis process.
- The weights of the samples were measured using weighing balance.
- Then the sample is filled inside the reactor and that can be properly sealed with the help of metallic gasket and bolt and nuts to make reactor leak free.
- The reactor is insulated fully with the wool and its' is supported with the help of ceramic.
- Outside pipe is connected to the condenser of counter flow type and it is further connected to the liquid separator.

The chipped material is filled in the reactor initially, and then reactor has closed with the help of bolt. Here gasket was used to prevent leakage. Then supply nitrogen gas from the cylinder to reactor for the time period of 3 to 4 min, after its closed. Switch on the electrical supply, initially set the temperature up to 600 ˚C in the temperature controller. For condenser water is supplied from inlet to outlet. Gases are collected in the balllon. Finally note the time taken to reach the temperature up to 600˚C. Cooling time of the reactor is 16 to 18 hours. Finally the pyrolysis oil, char and syngases are collected separately. This experiment is conducted with and without catalyst. In the catalytic pyrolysis catalyst is used in 1:10 ratio to feed mass. In temperature fractional pyrolysis is the oil is collected in 100 ˚C interval starting from 100 ˚C.

Table-1: Comparisons of Various Catalysts

Catalyst	Seed (kg)	Collected oil (ml)	Charcoal (gm)	Gas (gm)	Flash Point (˚c)	Fire point (˚c)
Without catalyst	1	250	350	400	41	45
Nacl	1	280	308	512	40	45
KOH	1	250	365	485	46	51
$CaCO_3$	1	260	355	485	44	48

Sources: Authors Analysis.

6. CALORIFIC VALUE

Table-2: Diesel Calorific Value is 44000 kJ / kg

S.No	Sample description	Calorific value (cal/g)
1	Nacl catalyst	1446
2	$CaCO_3$ catalyst	3561
3	KOH catalyst	1585

Sources: Authors Analysis.

7. RESULTS AND DISCUSSIONS

Three different experiments are conducted on pyrolysis of acacia arabica. These experiments are shown below:

- Total extraction of oil for 1kg of acacia arabica.
- Different yield temperature oil collection for 3kg of acacia arabica.
- Total extraction of oil for 1kg of acacia arabica using potassium hydroxide (kOH), sodium chloride (Nacl), calcium carbonate (CaCO3) as catalyst.

8. CONCLUSIONS

The catalyst like KOH, $CaCO_3$, Nacl is selected according to its oxidation characteristics and is added with acacia arabica. The pyrolysis process is carried out at a temperature of 620 - 700˚C. As a result of adding the catalyst, there is an increase in the extraction of oil. The increase in extraction of oil product quantity is compensated mainly with decrease in gas product. Of three catalysts, Nacl alone has the greatest power of the extraction of oil from biomass. The optimum oil product yield of 20% is achieved at 620˚C by using the Nacl.

9. RELATED LITERATURES

[1] Soo, Hyun Chung; Jong Park; Sang Goo Jeon, and Dong, Chan Kim, "Pyrolysis of Waste Plastic Using Synthesized Catalyst from Fly Ash", Korea Institute of Energy Research, 71-72, Jang-Dong, Yysong-gu, Taejon, Korea, 05-4.

[2] Yang, Yongrong; Chen, Jizhong, and Zhao, Guibin, "Technical Advances on the Pyrolysis of Used Tyres in China", Dept. of Chemical Engg., Zhejiang University Yuquan Campus, Hangzhou, 10027, P. R. CHINA.

[3] Zuzana, Krenova; Mohammad. Juma; Kristina. Holikova; Ludovit, Jelemensky, and Jozer, Markos, (2005). "Experimenal Study of Waste Rubber Pyrolysis and Combustion", Department of Chemical Na Biochemical Engineering. Petroleum & Coal 47(2), pp. 1-9.

[4] Laoy, Saeed; Anti, Tohka, and Ron, Zevenhoven, (2003). "Experimental Assessment of Two-Stage Fluidised Bed Combustion of High - PVC Solid Waste with HCI Recovery", Helisinki University of Technology, May, pp. 8-21.

[5] A., M. Fernandez; M., A. Diez, and R., Alvarez, and C., Barriocanal Institution del carbon (INCAR-CSIC). Apartado 73, 33080 Oviedo. Spain, "Pyrolysis of Tyre Wastes".

[6] A., Demirbas, (2001). "Energy Convers", Manage, 42, pp. 1357–1378.

[7] C., Yang; B., Zhang, and J., Moen, et al., (2010). International Journal of Agric. Biol. Eng. 3(3), pp. 54–61.

[8] M., Crocker, and C., Crofcheck, (2006). Energia, 17, pp. 6.

[9] K., Medlock, and R., Soligo, (2001), Energy J., 22 (2), pp. 77–106.

[10] J., G. Speight, (2011). Handbook of Industrial Hydrocarbon Processes, Elsevier Inc., ISBN: 978-0-7506-8632-7.

[11] Q., Zhang; J., Chang, and T., Wang, et al., (2006), Energy Fuels, 20, pp. 2717–2720.

[12] P., I. Borjesso, (1996). Energy Convers. Manage., 37, pp. 1235–1240.

[13] S., B. McLaughlin, and L., A. Kszos, (2005). Biomass Bioenergy, 28, pp. 515–535.

[14] S., B. Lee, and O., J. Fasina, (2009). Anal. Appl. Pyrolysis 86, pp. 39–43.

[15] S., B. McLaughlin; D., G., Ugarte, and C., T. Garten, et al., (2002). Environ. Sci. Technol., 36, pp. 2122–2129.

[16] A., M. Roth; D., W. Sample, and C., A. Ribic, et al., (2005). Biomass Bioenergy, 28, pp. 490–498.

[17] M., E. Walsh; D., G. Ugarte, and H., Shapouri, et al., (2003). Environ. Resour. Econ. 24, pp. 313–333.

[18] S., B. McLaughlin; J., Bouton, and B., Bransby, et al., (1999). Am. Soc. Hort. Sci. 20, pp. 282–299.

[19] P., R., Adler; M., A. Sanderson, and A., Akwasi, et al., (2006), Agron. J., 98, pp. 1518–1525.

[20] A., A. Boateng; K., B. Hicks, and K., P. Vogel, (2006). J. Anal. Appl. Pyrolysis, 75, pp. 55–64.

[21] R., Sage, and M., Russell, (1999). Plant Biol, 7, pp. 228-229, ISBN: 0126144400.

[22] N., Miskolczi; F., Buyong, and P., T. Williams, (2010). "Thermogravimetric Analysis and Pyrolysis Kinetic Study of Malaysian Refuse Derived Fuels", Journal of the Energy Institute, 83, pp. 25–132.

[23] H., Yang; R., Yan, and T., Chin, et al., (2004). Energy Fuels, 18, pp. 1814–1821.

[24] R., T. Rao, and A., Sharma, (1998). Energy, 23, pp. 973-978.

[25] M., Inguanzo; A., Domıˇınguez, and J., A. Meneˇındez, et al., (2002). J. Anal. Appl. Pyrolysis, 6, 209–222.

[26] J., A. Caballero; R., Font, and A., Marcilla, et al., (2001). J. Anal. Appl. Pyrolysis, 58, pp. 617–633.

[27] A., V. Bridgwater, (2000). "G.V.C. Peacocke", Renew. Sustain. Energy Rev., 4, pp. 1–73.

[28] Q., Liu; Z., Zhong; S., Wang, et al., (2010). J. Anal. Appl. Pyrolysis, 90, pp. 213–218.

[29] S., R., (2004). Czernik, A.V. Bridgwater, Energy Fuels, 18, pp. 590–598.

[30] H., Ukei; T., Hirose; S., Horikawa; Y., Takai; M., Taka; N., Azuma, and A., Ueno, "Catalytic Degradation of Polystyrene into Styrene and Design of Recyclable Polystyrene, with Dispersed Catalysts", Catalysis Today. 62. pp. 67–75.

[31] N., Miskolczi; F., Buyong; A., Angyal; P., T. Williams, and L., Bartha, (2010). "Two Stages Catalytic Pyrolysis of Refuse Derived Fuel: Production of Biofuel via Syncrude", Bioresource Technology, 101, pp. 8881–8890.

[32] D., P. Serrano; J., Aguado; J., and M., Escola, (2000). "Catalytic Conversion of Polystyrene Over HMCM-41, HZSM-5 and Amorphous SiO2–Al2O3: Comparison with Thermalcracking", Applied Catalysis B: Environmental, 25, pp. 181–189.

[33] F., Pinto; P., Costa; I., Gulyurtulu, and I. Cabrita, (1999). "Pyrolysis of Plastic Waste 2 Effect of Catalyst on Product Yield", Journal of Analytical and Applied Pyrolysis, 51, pp. 57–71.

[34] Y., C. Liang; C., Y. May; C., S. Foon; M., A. Ngan; C., C. Hock, and Y., Basiron, (2006). "The Effect of Natural and Synthetic Antioxidants on Oxidative Stability of Palm Diesel", Fuel, 85, pp. 867–870.

[35] G., Knothe, (2007). "Some Aspects of Biodiesel Oxidative Stability", Fuel Processing Technology, 88, pp. 669–677.

[36] A., M. Missaoui; V., A. Fasoula, and J. H. Bouton, (2005). Euphytica, 142 (142), pp. 1–12.

*****

EFFICIENT BROADCASTING AND ROUTING TO MINIMIZE TRANSMISSION DELAY FOR CRITICAL EVENT MONITORING

Krishna Priya C. B.[32] N. Mookhambika[33]

ABSTRACT

Sensor nodes are expected to work for a long time for event monitoring without recharging their batteries. So sleep scheduling method is always used during the monitoring process. When a critical event occurs, an alarm message should broadcast to the entire network as soon as possible. But it cause transmission delay because sender nodes should wait until receiver nodes are active and ready to receive the message.

The Existing Sleep scheduling methods are employed to prolong the network life time which in turn broadcast the alarm message to the entire network within stipulated time. Thus delay-efficient sleep scheduling method needs to be designed to ensure low broadcasting delay from any node in the WSN. It is designed with two determined traffic paths such as Path for transmission of alarm message and Level-by-level offset based wake-up pattern according to the selected path. The topology influences the performance of the network in terms of broadcasting, scheduling of transmission and routing. When a critical event occurs, an alarm message should broadcast to the entire network as soon as possible.

If such topology construction is poorly designed it causes transmissions delay. Although the existing system avoids delay there are still few challenges. So the proposed topology of a wireless network should be efficiently designed to achieve high network performance.

Thus remove useless topology information such as collision, redundancy and transmission delay from a network to dramatically improve a network's broadband utilization, delivery ratio, network lifetime and packet retransmission. This is called as topology control management. Thus Connected Dominating Set (CDS) is regarded as an effective approach to hierarchical topology organization.

KEYWORDS

Sensor, Connected Dominating Set, Network, Nodes, Scheduling, Wireless, Broadcast etc.

INTRODUCTION

In mission-critical applications, such as battlefield reconnaissance, fire detection in forests, and gas monitoring in coal mines, wireless sensor networks (WSNs) are deployed in a wide range of areas, with a large number of sensor nodes detecting and reporting some information of urgencies to the end-users. As there may be no communication infrastructure, users are usually equipped with communicating devices to communicate with sensor nodes.

When a critical event (e.g., gas leak or fire) occurs in the monitoring area and is detected by a sensor node, an alarm needs to be broadcast to the other nodes as soon as possible, which is shown in Fig. 1 as an example. Then, sensor nodes can warn users nearby to flee or take some response to the event. As sensor nodes for event monitoring are expected to work for a long time without recharging their batteries, sleep scheduling method is always used during the monitoring process. Obviously, sleep scheduling could cause transmission delay because sender nodes should wait until receiver nodes are active and ready to receive the message. The delay could be significant as the network scale increases. Therefore, a delay-efficient sleep scheduling method needs to be designed to ensure low broadcasting delay from any node in the WSN.

To minimize the broadcasting delay, it is needed to minimize the time wasted for waiting during the broadcasting. The ideal scenario is the destination nodes wake up immediately when the source nodes obtain the broadcasting packets. Here, the broadcasting delay is definitely minimum.

However, it is still a challenge for us to apply the level-by-level offset to alarm broadcasting in the critical event monitoring. First, the order of nodes' wake-up should conform to the traffic direction. If the traffic flow is in the reverse direction, the delay in each hop will be as large as the length of the whole duty cycle. Second, the level-by-level offset employed by the packet broadcasting could cause a serious collision. Finally, the transmission failure due to some unreliable wireless links may cause the retransmission during the next duty cycle, which also results in large delay equaling the whole duty cycle.

The novel sleep scheduling method which is still based on the level-by-level offset schedule to achieve low broadcasting delay in a large scale WSN. As the alarm message may be originated by any possible node, we set two phases for the alarm broadcasting in the proposed sleep scheduling method.

First, when a node detects a critical event, it originates an alarm message and quickly transmits it to a center node along a predetermined path with a level-by-level offset way. Then, the center node broadcasts the alarm message to the other nodes along another path also with a level-by-level offset way. Through designing a special wake-up pattern, the two possible traffics could be both carried by a node and the node just needs to be awake for no more than time in each duty cycle to transmit an alarm packet.

For the critical event monitoring in a WSN, sensor nodes are usually equipped with passive event detection capabilities that allow a node to detect an event even when its wireless communication module is in sleep mode. Upon the detection of an event by the sensor, the radio module of the sensor node is immediately woken up and is ready to send an alarm message

APPLICATIONS OF DOMAIN

- In Public WLAN provides campus-wide indoor and outdoor coverage.
- It provides flexible solution to implement the information delivery system required to control transportation services.

[32] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mayjamv@gmail.com
[33] Assitant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

- Wildlife monitoring focuses on tracking wild species to deeply investigate their behaviour and understand the interactions and influences on each other, as well as their reaction to the ecosystem changes caused by human activities.

- Opportunistic networks can provide intermittent Internet connectivity to rural and developing areas where they typically represent the only affordable way to help bridging the digital divide.

- VANETs use ad hoc communications for performing efficient driver assistance and car safety. The communications include data from the roadside and from other cars. VANET research aims to supply drivers with information regarding obstacles on the road and emergency events, mainly due to line-of-sight limitations and large processing delays. VANET can be used to communicate premonitions, notification of emergencies, and warnings about traffic conditions.

- The underwater wireless sensor network have applications including the scientific (e.g., oceanographic data collection for scientific exploration, pollution control, or climate monitoring), military (e.g., tactical surveillance), and civilian fields (e.g., tsunami warnings).

PROBLEM STATEMENT

To reduce the energy consumption of overhearing, nodes not involved in current transmission should shut down their transceivers. Ideally in a sensor network, a scheduling protocol must determine a transmission schedule for each packet to avoid collisions. Such fine-grained scheduling can be performed centrally or distributed. For central computation the overhead of control messages and the delay in scheduling is often prohibitive.

Although an out-of-band channel and an extra radio could be used to facilitate the scheduling, the additional cost, and the added complexity of managing the power in the second radio makes it a less attractive alternative. Distributed computation of fine-grain scheduling is also faced with problems of extra messaging, conservative channel assignment and thus low channel utilization.

Sensor nodes are expected to work for a long time for event monitoring without recharging their batteries. So sleep scheduling method is always used during the monitoring process. When a critical event occurs, an alarm message should broadcast to the entire network as soon as possible. But it cause transmission delay because sender nodes should wait until receiver nodes are active and ready to receive the message

OBJECTIVES OF STUDY

Sleep scheduling methods are employed to prolong the network life time which in turn broadcast the alarm message to the entire network within stipulated time. Thus delay-efficient sleep scheduling method needs to be designed to ensure low broadcasting delay from any node in the WSN. The objective is to reduce the delay of alarm broadcasting from any sensor node in WSNs.

It is challenging to achieve energy-efficient, low latency and reliable multihop broadcast over asynchronous dutycycling. For example, the neighbors of a transmitter wake up asynchronously, requiring the transmitter to stay active long enough so that each neighbor has a chance to receive the broadcast packet, resulting in increased energy consumption.

In addition, transmission attempts over poor quality links can significantly decrease delivery ratio and increase delivery latency and energy consumption. When a transmission fails and the intended receiver goes to sleep, if the transmitter is to retransmit, it must wait until the receiver wakes up in next cycle. A transmitter may also substantially delay forwarding by other neighbours, if the transmitter occupies the medium while waiting to reach all of its neighbors.

Finally, information about the progress of a broadcast is important for a node to avoid redundant transmissions, but a node that has just taken up has no up-to-date progress information. The node cannot simply use overhearing to learn this information as the progress may change while the node has its radio off.

A novel sleep scheduling method in the existing system is designed with two determined traffic paths. They are:

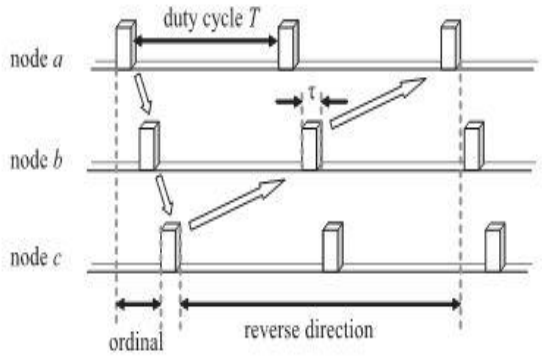
- A Path for transmission of alarm message.

- Level-by-level offset based wake-up pattern according to the selected path.

When a critical event occurs, an alarm is quickly transmitted along one of the traffic paths to a center node and then it is immediately broadcast by the center node along another path without collision. Thus broadcasting delay is independent of the density of nodes and its energy consumption is ultra low. Exactly, the upper bound of the broadcasting delay is only 3D + 2L, where D is the maximum hop of nodes to the center node, L is the length of sleeping duty cycle.

The packet can be delivered from **node a to node c** via node b with minimum delay. Hence it is possible to achieve low transmission delay with the level-by-level offset schedule in multi-hop WSNs. But sending alarm broadcasting in the critical event monitoring is difficult.

Figure-1: The level-by-level Offset Schedule.



Sources: Authors Compilation.

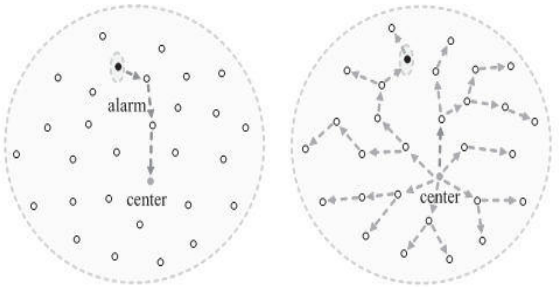
Issues in level-by-level offset to alarm in the critical event monitoring:

1. The order of nodes' wake-up should conform to the traffic direction. If the traffic flow is in the reverse direction the delay in each hop will be as large as the length of the whole duty cycle.

2. The level-by-level offset could cause transmission failure due to some unreliable wireless links. So retransmission during the next duty cycle results in delay.

It is known that the alarm could be originated by any node which detects a critical event in the WSN. To essentially reduce the broadcasting delay, includes two phases:

Figure-2: Illustrates Two Phases of Processing.



Sources: Authors Compilation.

1) Any node which detects a critical event sends an alarm packet to the center node along a predetermined path according to level-by-level offset schedule.

2) The center node broadcasts the alarm packet to the entire network also according to level-by-level offset schedule.

Minimize the Broadcast Delay- special wake up pattern

In the scheduling scheme each node needs to wake up properly for both of the traffics in two parts:

1) Establish the two traffic paths in WSN:

   - The traffic path from node to the center node is defined as uplink- breadth first search (BFS) tree.

   - The traffic path from the center node to other nodes is defined as downlink- colored connected dominant set via IMC algorithm is established. Each node transmits or receives packets in a specific channel according to the color assigned. Thus eliminate the collision in broadcasting.

2) Calculate the wake-up parameters (e.g., time slot and channel) for all nodes to handle all possible traffics.

DISADVANTAGES

- If size of time slot is enlarged, the performance of the existing scheme could not be further encouraged.

- Enlargement of time slot increases energy consumption of sensor nodes. This is because nodes have to keep awake during the whole of the synchronous time slot.

The topology of a wireless network should be efficiently designed to achieve high network performance. Thus remove useless topology information such as collision, redundancy and transmission delay from a network to dramatically improve a network's broadband utilization, delivery ratio, network lifetime and packet retransmission. This is called as topology control management.

Topology control is the efficient technique in reducing network transmission power. The issues for topology control are stated with many reasons such as dominance of lossy and asymmetric links in WSNs. Moreover existing topology control schemes aim at maintaining connectivity based network properties. However connectivity alone does not be adequate to provide communication performance when the network is lossy.

Communication along lossy network path may result in excessive packet loss and energy waste. Therefore topology control must minimize the power consumption of the network while achieving the desired path quality. Hence the issue of link unreliability is resolved using the proposed new topology control metrics known as localized configurable topology control (CTC) algorithm.

This algorithm provides path quality assurance over lossy and asymmetric links in WSNs. Further it enables applications to achieve desired trade off between transmission power and path quality based on their specific requirements which handle dynamic network efficiently.

Topology control includes power control and hierarchical topology organization. The target of power control is to adjust nodes' transmission range to achieve balanced connectivity while hierarchical topology organization aims to find a communication backbone from the original network in charge of all forwarding's in the network.

The proposed system considers both per node and per-link power control strategies. While per-node control assigns each node a single power, per-link control may assign a node different power for different links originating from it. The two optimization metrics is known as min_sum and min_max metric. The min_sum minimizes the total power of all nodes or links in the network and min_max minimizes the maximum power among all nodes or links. The min_max metric may lead to a longer network lifetime by balancing the power consumption of different nodes.

CTC uses a two-hop neighborhood graph that is constructed from link quality information. Node v is node u's one-hop neighbor if there exists at least one link (u,v, i) from u to v. The one-hop neighbourhood graph of u includes u and all the one-hop neighbours of u and all the links from u to its neighbors.

The twohop neighborhood graph of node u is the union of the one-hop neighborhood graphs of u and u's neighbors. We use $Ni(u) = (Vi(u),Ei(u))$ $(i = 1,2)$ to denote the one-hop and two-hop neighborhood graphs at u. Although links may be asymmetric, we require neighborhood relation to be symmetric. Each node u can enforce this requirement by pruning the links to the neighbors who do not include u within their one-hop

neighborhood. In order to construct N2(u), node u needs to collect the transmission counts of the links within its two-hop neighbourhood at different power levels. Each node can measure transmission counts of its one-hop links based on data or hello messages, and exchange them with its one-hop neighbors.

- Efficient routing cost influences the performance of the network in terms of broadcasting, scheduling of transmission and routing.

- Designed high efficient topology construction reduces transmissions delay at times of critical event which broadcast to the entire network as soon as possible. (MIS) in G; 2) select connector nodes to form a connected dominated set (CDS), and partition connector nodes and independent nodes in each layer into four disjoint sets with IMC algorithm. All nodes have been divided into H1, H2, H3; HD with the BFS tree, the MIS can be established layer by layer (i.e., hop by hop) in the BFS as follows: Start from the 0th hop, we pick up a maximum independent set, then, move on to the first hop, pick up another maximum independent set. Note that, independent nodes of the first hop also need to be independent of those in the previous hop. Repeat this process until all hops of nodes have been worked on.

  Second, we construct the CDS by selecting connector nodes C from V nI to interconnect independent nodes as follows: Obviously, for any two 2-hop neighboring independent nodes, at least one node in G is adjacent to both of them. Hence, the node is possible to be selected as a connector node.

WAKE-UP PATTERNS

After all nodes get the traffic paths, sending channels and receiving channels with the BFS and CCDS, the proposed wake-up pattern is needed for sensor nodes to wake-up and receive alarm packet to achieve the minimum delay for both of the two traffic paths. As described above, there are two traffic paths for the alarm dissemination, and sensor nodes take two level-bylevel offset schedules for the traffic paths. In level-by-level offset schedules: 1) sensor nodes on paths in the BFS wake up level-by-level according to their hop distances to the center node; and 2) after the center node wakes up, the nodes in the CCDS will go on to wake up level-by-level according to their hop distances in the CCDS.

Hence, when an alarm packet is originated, it could be quickly forwarded to the center node along a path in the BFS, then, the center node immediately broadcasts it along the paths in the CCDS. Since it is hard to predict when the alarm occurs, the two level-by-level offset schedules are taken periodically moreover, it is needed to effectively arrange time slots for sensor nodes at different positions in the topology, so that two level-by-level offset schedules can periodically work without interfering with each other.

MULTIPLE ALARMS

In some cases, the critical event may trigger several alarms in the network, and they may be sent to a parent node when it wakes up. To deal with the collision, we design a mechanism for the proposed scheduling as follows: When a sensor node having detected the event is going to send an alarm packet, it keeps transmitting the packet randomly with the probability during the time slot.

However, if the node detects some others are transmitting alarm packets during the same time slot, it gives up its transmission. Through this way, the nodes sending alarms could be decreased gradually.

The parent node cannot judge whether there is an alarm packet by just detecting the channel, because some configuration packets also need to be transmitted in the network and the alarm packet needs to be exactly received to avoid misinformation

LOCALISED TOPOLOGY CONTROL ALGORITHM

The objective is to choose a power assignment Ω with the minimum sum while the DTC bound of the induced topology under Ω is no greater than t:

$$\Omega = argmin \sum_{P_i \in \Omega} P_i, subject\ to \max_{u,v \in V} \frac{\Gamma_{G_\Omega}(u,v)}{\Gamma_{G_M}(u,v)} \leqslant t$$

Ω include transmission power for each node or link depending whether the per-node control or per-link control strategy is used. G_X (u, v) denotes the minimum transmission count from u to v in the network under power assignment.

G_M denotes the topology where each node is assigned the maximum power. G_M achieves the best path quality among all topologies under any possible power assignment when the network workload is light. G_Ω represent topology induced by the power assignment Ω. We define the dilation of transmission count (DTC) of G_Ω as the maximum ratio of the minimum transmission count between any two nodes in G_Ω to that between the same nodes in G_M. DTC quantifies the worst-case degradation in network's path quality under a power assignment relative to the maximum-power case. This metric closely relates to communication performance like reliability, throughput and delay.

PERFORMANCE EVOLUTION

Before applying the proposed scheme, some initialization works need to be done first, such as obtaining the topology and broadcasting the assignment. We evaluate the performance of the mechanism with a simple and typical network model.

Suppose there are M nodes that need to send packets to a parent node which keeps awake for 20 ms every two duty cycles periodically. The quality of the link between the parent node and each child is 70 percent. It shows the time when the parent node successfully receives a packet.

CONCLUSIONS AND FUTURE WORK

The sleep schedule scheme could essentially decrease the delay of alarm broadcasting from any node in WSN. Moreover, the alarm broadcasting delay is independent of the density of nodes in WSN. The topology influences the performance of the network in terms of broadcasting, scheduling of transmission

and routing. When a critical event occurs, an alarm message should broadcast to the entire network as soon as possible. If such topology construction is poorly designed causes transmissions delay. To achieve efficient routing, minimum routing cost preserving CDS is proposed. Hence we can achieve efficient routing through CDS.

Thus this project incorporates lossy link models in the design of topology control algorithms using CTC for WSNs. The capability of configuring the topology of a lossy WSN achieves desired path quality bounds in a localized fashion.

It provides desired trade off between power consumption and network performance according to application requirements. The efficiency is achieved in terms of both communication performance and energy consumption

REFERENCES

[1] N., Bouabdallah; M., E. Rivero-Angeles, and B. Sericola, (2009). "Continuous Monitoring Using Event-Driven Reporting for Cluster-Based Wireless Sensor Networks", IEEE Trans. Vehicular Technology, September, Volume: 58, Number: 7, pp. 3460-3479.

[2] M., I. Brownfield; K., Mehrjoo; A., S. Fayez, and N., J. Davis IV., (2006). "Wireless Sensor Network Energy-Adaptive Mac Protocol", Proceeding Third IEEE Consumer Comm. and Networking Conf., January, pp. 778-782.

[3] T., Zheng; S., Radhakrishnan, and V., Sarangan, (2005). "PMAC: An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks", Proceedings 19th IEEE International Parallel and Distributed Processing Symp., April, pp. 224-231.

[4] S., C. Ergen, and P., Varaiya, (2010). "TDMA Scheduling Algorithms for Wireless Sensor Networks", Wireless Networks, Volume: 16, Number: 4, pp. 985-997.

[5] G., Lu; B., Krishnamachari, and C., Raghavendra, (2004). "An Adaptive Energy-Efficient and Low-Latency MAC for Data Gathering in Wireless Sensor Networks", Proceeding 18th IEEE International Parallel and Distributed Processing Symp., April, pp. 224-230.

[6] A., Keshavarzian; H., Lee, and L., Venkatraman, (2006). "Wakeup Scheduling in Wireless Sensor Networks", Proceeding Seventh ACM International Conference Mobile Ad Hoc Networking and Computing, May, pp. 322-333.

[7] G., Lu; N., Sadagopan; B., Krishnamachari, and A., Goel, (2005). "Delay Efficient Sleep Scheduling in Wireless Sensor Networks", Proceedings 24th IEEE International Conference Computer Comm., March, pp. 2470-2481.

[8] N., A. Vasanthi, and S., Annadurai, (2006). "Energy Efficient Sleep Schedule for Achieving Minimum Latency in Query Based Sensor Networks", Proceeding IEEE International Conference Sensor Networks, Ubiquitous, and Trustworthy Computing, June, pp. 214-219.

[9] N., A. Vasanthi, and S., Annadurai, (2006). "AWS: Asynchronous Wakeup Schedule to Minimize Latency in Wireless Sensor Networks", Proceeding IEEE International Conference Sensor Networks, Ubiquitous, and Trustworthy Computing, June, pp. 144-151.

[10] Y., Sun; O., Gurewitz; S., Du; L., Tang, and D., B. Johnson, (2009). "ADB: An Efficient Multihop Broadcast Protocol Based on Asynchronous Duty-Cycling in Wireless Sensor Networks", Proceeding Seventh ACM Conference Embedded Networked Sensor Systems, November, pp. 43-56.

*****

WINEAR FILTER BASED IMAGE DENOISING

Shanmuga Priya N.[34] S. Sangeetha[35]

ABSTRACT

We propose a denoising method motivated by our previous analysis of the performance bounds for image denoising. Insights from that study are used here to derive a high-performance practical denoising algorithm. We propose a patch-based Wiener filter that exploits patch redundancy for image denoising. Our framework uses both geometrically and photo-metrically similar patches to estimate the different filter parameters. We describe how these parameters can be accurately estimated directly from the input noisy image.

KEYWORDS

Image, Denoising, Algorithm, Patches etc.

RELATED WORK

We derived an expression for the fundamental limit to image denoising assuming that the noise-free image is available. Estimator for bound on the mean squared error gives only noisy image and noise characteristics. To do this, we make use of an assortment of independently collected noise-free images from which prior information about the noisy image is learned. We show that even for reasonably low input signal-to-noise levels, our method can predict the denoising bound with accuracy. The challenge of any image denoising algorithm is to suppress noise while producing sharp images without loss of finer details. The first modern adaptive method to successfully address these contradictory goals can be attributed to Tomasi et al. [4], where the authors proposed a generalization of the SUSANfilter [5], which itself was an extension of the Yaroslavky filter[6]. The authors there proposed denoising by weighted aver-aging pixels similar in intensity within a local neighborhood. Under strong noise, identifying such similar pixels can be challenging. In [7], Takeda et al. proposed a signal-dependentsteering kernelregression (SKR) framework for denoising. This method proved to be much more robust under strongnoise. A patch-based generalization of the bilateral filter [4] was proposed in [8] and [9], where the concept of localitywas extended to the entire image. Although the results there were encouraging, the true potential for this nonlocal means (NLM) method was only realized in [10] and [11].

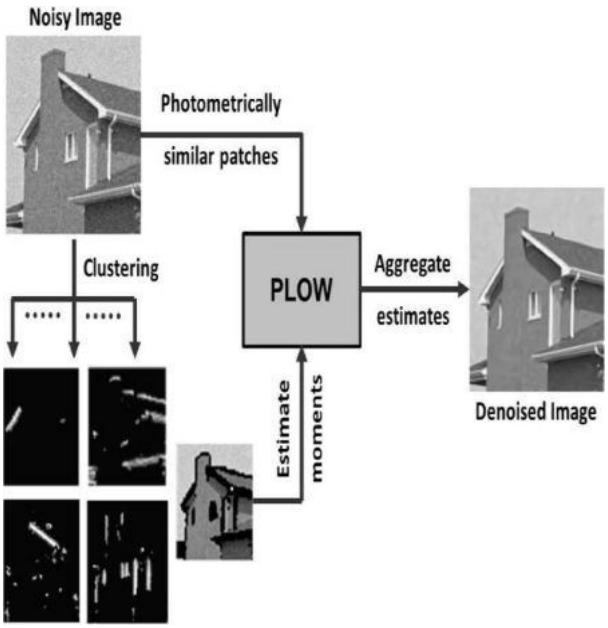
Another patch redundancy-based framework, i.e., BM3D [2], adopts a hybrid approach of grouping similar patches and performing collaborative filtering in some transform [e.g., discrete cosine transform (DCT)] domain. It ranks among the best performing methods that define the current state of the art.

PROPOSED SYSTEM

Our denoising approach, designed for near-optimal performance (in the mean-squared error sense), has a sound statistical foundation that is analyzed in detail. The performance of our approach is experimentally verified on a variety of images and noise levels.

The results presented here demonstrate that our proposed method is on par or exceeding the current state of the art, both visually and quantitatively. We first identify geometrically similar patches within the noisy image. Once such patches are identified from its noisy observations, such features need to be robust to the presence of noise, as well as to differences in contrast and intensity among patches exhibiting similar structural characteristics. Possible choices of features include contrast-adjusted image patches [3] or principal components in conjunction with predetermined clustering guides [2]. For our purposes, where the image patches can be considerably noisy, we make use of the locally adaptive regression kernels (LARKs) introduced for denoising in [7] and subsequently adapted as features for geometric clustering [5] and object detection [3]. We refer the interested reader to [7], where the design of the kernels is covered in detail.

Figure-1



Sources: Authors Compilation.

CONCLUSIONS

We have proposed a method of denoising motivated from our previous work in analyzing the performance bounds of patch-based denoising methods. We have developed a locally optimal Wiener-filter-based method and have extended it to take advantage of patch redundancy to improve the denoising performance. Our denoising approach does not require parameter tuning and is practical, with the added benefit of a clean statistical motivation and analytical formulation. We analyzed the framework in depth to show its relation to nonlocal means and residual filtering methods such as, Through Experimental Validation, we have shown that our method produces results quite comparable with the state of art.

[34] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, priyashacse@gmail.com
[35] Assitant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, ssangi05@gmail.com

REFERENCES

[1] P., Chatterjee, and P., Milanfar, (2010). "Is Denoising Dead?", *IEEE Trans. Image Process.*, April, Volume: 19, Number: 4, pp. 895–911.

[2] P., Chatterjee, and P., Milanfar, (2011). "Practical Bounds on Image Denoising: From Estimation to Information", *IEEE Trans. Image Process.*, May, Volume. 20, Number; 5, pp. 1221–1233.

[3] G., E. Healey, and R., Kondepudy, (1994). "Radiometric CCD Camera Calibration and Noise Estimation", *IEEE Trans. Pattern Anal. Mach. Intell.*, March, Volume: 16, Number: 3, pp. 267–276.

[4] C., Tomasi, and R., Manduchi, (1998). "Bilateral Filtering for Gray and Color Images", in *Proceeding International Conference Comput. Vis.*,Washington, DC, January, pp. 839–846.

[5] S., M. Smith, and J., M. Brady, (1997). "SUSAN—A New Approach to Low Level Image Processing", *International Journal Comput. Vis.*, May, Volume: 23, Number: 1, pp. 45–78.

[6] L., P. Yaroslavsky, *(1985). Digital Picture Processing,* Secaucus, NJ: Springer-Verlag.

[7] H., Takeda; S., Farsiu, and P., Milanfar, (2007). "Kernel Regression for Image Processing and Reconstruction", *IEEE Trans. Image Process.*, February, Volume: 16, Number: 2, pp. 349–366.

[8] A., Buades; B., Coll, and J., M. Morel, (2005). "A Review of Image Denoising Methods, with a New One", *Multiscale Model. Simul.*, Volume: 4, Number: 2, pp. 490–530.

[9] S., P. Awate, and R., T. Whitaker, (2006). "Unsupervised, Information-Theoretic, Adaptive Image Filtering for Image Restoration", *IEEE Trans. Pattern Anal. Mach. Intell.*, March, Volume: 28, Number: 3, pp. 364–376.

[10] C., Kervrann, and J., Boulanger, (2006). "Optimal Spatial Adaptation for Patchbased Image Denoising", *IEEE Trans. Image Process.*, October, Volume: 15, Number: 10, pp. 2866–2878.

[11] C., Kervrann, and J., Boulanger, (2008). "Local Adaptivity to Variable Smoothness for Exemplar-Based Image Denoising and Representation", *International Journal Comput. Vis.*, August, Volume: 79, Number: 1, pp. 45–69.

****

SURGICAL INSTRUMENT TRACKING USING RFID

R. Surya Dharshini[36] N. Mookhambika[37]

ABSTRACT

Our paper aims to develop an electronic inventory that can track every surgical instrument used during surgery and also it checks for sterilization using RFID. Every surgical instrument used during surgery must be accounted for after surgery to ensure that none of these is left inside the patient. Despite the numerous precautions in place, in approximately 1 in 1500 cases, something gets left behind inside the patient's body.

This paper aims at increasing the safety of surgical procedures. This surgical instrument tracking system utilizes RFID (Radio Frequency Identification) technology along with mapping and localization to aid in accounting for all instruments used during surgery. RFID is a unique serial number for each instrument that can be received by wireless means.

KEYWORDS

RFID, Surgical, Sterilization, Technology etc.

RADIO FREQUENCY IDENTIFICATION (RFID)

WORKING & PRINCIPLE OPERATION

Radio frequency identification is so-named because it relates to the identification of objects using EM (Electro magnetic) radiation at radio frequencies. RFID systems may be categorized based on the band of the EM spectrum that they operate in.

RFID systems in the same band will generally display similar characteristics; those in other bands may well operate very differently and therefore by more or less suitable for a given application. An RFID system comprises two components – an RFID reader and an RFID code module. Despite its name, the RFID code module is really the transmitter in an RFID system.

The electronics in the code module uses an external power source to generate the signal that drives code module's antenna and which in turn creates the appropriate radio wave. This radio wave may be received by the RFID reader, which in turn 'reflects' some of the energy it receives in a particular way. Whilst this reflection is going on, the RFID reader is also acting as a radio receiver, so it can detect and decode the reflected signal in order to identify code.

STERILIZATION CHECK

For sterilization check, all the instruments are placed on a conveyor belt. The RFID reader is also placed near the conveyor belt set up. The database of sterilized instruments from the Central Sterilization and Supply Department (CSSD) is maintained in a computer in the operation theatre.

[36]Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, suryadharshini.raman@gmail.com
[37] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

The receiver in the RFID reader receives the RFID code from the instrument and decodes it and then sends the signals to the microcontroller. The Microcontroller is programmed in such a way to analyze the signal from reader and data from computer. The instruments that are sterilized and not sterilized are displayed in the computer.

TRACKING INSIDE PATIENT BODY

Once the surgery is completed, the RFID reader is activated to read if any instrument is placed inside the patient's body. If any instrument is detected by the reader, the microcontroller receives the signal and it activates alarm.

The microcontroller is programmed to display the instrument's code in a Liquid Crystal Display unit and also in the computer kept in the operation theatre.

MAPPING AND LOCALIZATION

Here the operation theatre is divided into 4 zones. The RFID readers when activated locate the instruments placed in the operation theatre and send the data to the microcontroller. Thus an instrument placed anywhere in the operation theatre can be located by receiving and decoding the codes by the reader and it can be displayed in the computer in which particular zone the instrument is found.

Decoder

The input is given to the pin 14 from the receiver. This data is decoded and can be collected at pin 10-13. A0-A7 is the address pins using this signal collision are avoided. VDD is connected with the supply. The supply is noise filtered by capacitors. Diode prevents encoder from wrong connection. Resistor across the osc1 and osc2 supports oscillator.

Encoder

In this circuit HT12E encoder IC is used to encode the data. The inputs are given through the pins 10-13. This data is encoded and can be collected at pin 17. It is given to the transmitter data input pin no. 2A0-A7 is the address pins and AD8-AD11 is the address/data pins' using this signal collision is avoided. When the transmission enable signal is low data AD8-AD11 are acts as a data pins and give the data to data out. VDD is connected with the supply. When TE is get high the encoder stops the cycle and Dout is stopped. The supply is noise filtered by capacitors. Diode prevents encoder from wrong connection. Resistor across the osc1 and osc2 supports oscillator.

Microcontroller

The AT89S52 is a low-power, high-performance CMOS 8-bit microcontroller with 8Kbytes of in-system programmable flash memory. The device is manufactured using Atmel's high-density nonvolatile memory technology and is compatible with the industry-standard 80C51 instruction set and pin out.

The on-chip 32 .Flash allows the program memory to be reprogrammed in-system or by a conventional nonvolatile memory programmer. By combining a versatile 8-bit CPU with

in-system programmable Flash on a monolithic chip, the Atmel AT89S52 is a powerful microcontroller which provides a highly-flexible and cost-effective solution to many embedded control applications.

The AT89S52 provides the following standard features: 8K bytes of Flash, 256 bytes of RAM, 32 I/O lines, Watchdog timer, two data pointers, three 16-bit timer/counters, a six-vector two-level interrupt architecture, a full duplex serial port, on-chip oscillator, and clock circuitry. In addition, the AT89S52 is designed with static logic for operation down to zero frequency and supports two software selectable power saving modes.

The Idle Mode stops the CPU while allowing the RAM, timer/counter, serial port, and interrupts system to continue functioning. The Power-down mode saves the RAM contents but freezes the oscillator, disabling all other chip functions until the next interrupt or hardware reset.

The major Features of 8-bit Micro controller **ATMEL 89S52:**

- 8K Bytes of In-System Programmable (ISP) Flash Memory.
- 4.0V to 5.5V Operating Range.
- Fully Static Operation: 0 Hz to 33MHz.
- Three-level Program Memory Lock.
- 256 x 8-bit Internal RAM.
- Three 16-Bit Timer/Counters.
- 32 Programmable I/O Lines.
- Eight Interrupt Sources.
- Full Duplex UART Serial Channel.
- Low-power Idle and Power-down Modes.
- Interrupt Recovery from Power-down Mode.
- Watch Dog Timer.

SOFTWARE: ALGORITHM

STEP1: MUC and COM ports are initialized.

STEP2: The codes' from various instruments are read and is fed to the MUC.

STEP3: Sterilization Check

The MUC cross checks the codes with that of the PC and is displayed. Alarm is activated and simultaneously it is displayed.

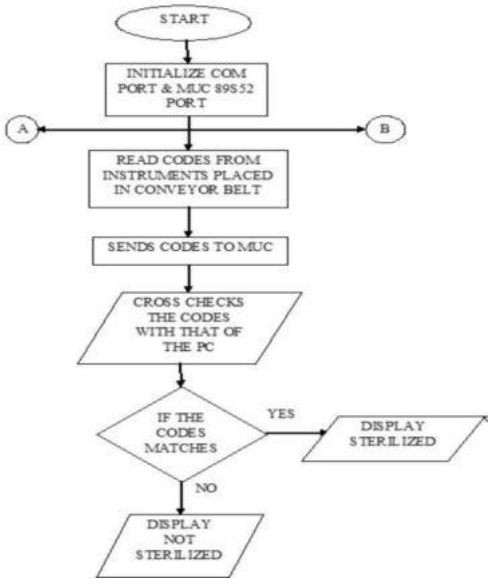
STEP4: Tracking Inside Patient Body

If any instrument is present inside the patient body, the reader detects it and alarm is activated and simultaneously it is displayed. .

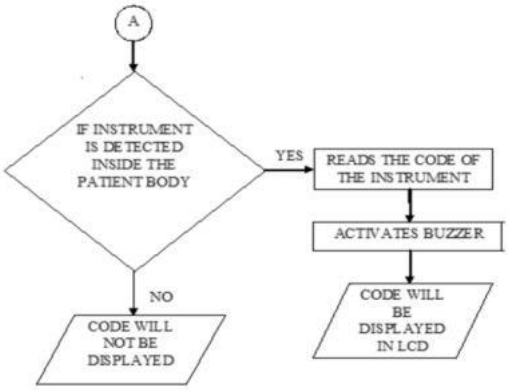
STEP5: Mapping and Localization

When the reader is activated, if any instrument is present, in the OT room, it will be displayed in PC.

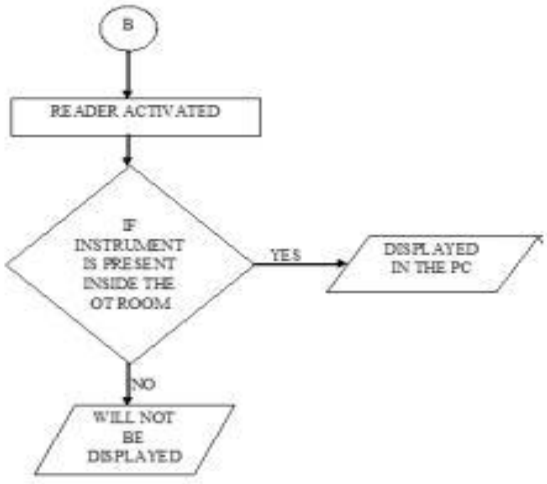
FLOWCHART



FLOWCHART OF A:



FLOWCHART OF B:



VISUAL BASIC

Advanced features in Visual Basic 5.0 such as optimizing native code compilation, accelerated from rendering, and enhanced database access allow developers to create fast, high performance applications and components.

Add to these features the new customizable development environment with IntelliSense technology, and developers will work with even greater productivity. In our project VB coding is utilized to display sterilization check, tracking inside the patient body, mapping and localization This include the output screen of the display of the various features being done in our project.

CONCLUSIONS

At present, in our paper we have used four instruments attached with an encoder circuit. In future, we can miniaturize the circuit, sterilization check can be done for many instruments within a short time as it is computerized and using Global Positioning System (GPS) we can trace out instrument anywhere inside the hospital. Hence we can ensure sterilization and assurance for patient safety.

REFERENCES

[1] Roy, Chowdhary, (2003). "Linear Integrated Circuits", New Age International (P) Ltd Publishers, Second Edition: 2003, Reprint: 2004.

[2] Krishna, S. Jaya., "RFID: Applications and Cases".

[3] Smitha, Ramachandran, "Radio Frequency Identification: Concepts and Cases".

[4] W., Alan Davis, "Radio Frequency Circuit Design".

[5] (1999). "Recommended Practices for Sponge, Sharp, and Instrument Counts, Aorn Recommended Practices Committee. Association of Perioperative Registered Nurses", AORN J, Volume: 70, Number: 6, pp. 1083–9.

[6] C., Kaiser; S., Friedman; K., Spurling; T., Slowick, and H., Kaiser, (1996). "The Retained Surgical Sponge", Ann Surg, Volume: 224, Number: 1, pp. 79–84.

[7] A., Gonzalez-Ojeda; D., Rodriguez-Alcantar; H., Arenas-Marquez; E., Sanchez Perez-Verdia; R., Chavez-Perez; R., Alvarez-Quintero, and A., Perea-Sanchez, "Retained Foreign Bodies Following Intra-abdominal Surgery", Hepatogastroenterology, Volume: 46, Number: 26, pp. 808–12.

[8] C., Zhan, and M., Miller, (2003). "Excess Length of Stay, Charges, and Mortality Attributable to Medical Injuries During Hospitalization", JAMA, Volume: 290, Number: 14, pp. 1868–74.

[9] A., Gawande; D., Studdert; E., Orav; T., Brennan, and M., Zinner, (2003). "Risk Factors for Retained Instruments and Sponges after Surgery", N Engl J Med, Volume: 348, Number: 3, pp. 229–35.

[10] K., Finkenzeller, (2003). RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, 2nd ed.

[11] M., Levine; B., Adiba; K., Mandl; I., Kohane, and J., Halamka, (2007). "What Are the Benefits and Risks of fitting Patients with Radiofrequency Identification Devices", PLoS Medicine, Volume: 4, Number: 11.

[12] S., Sarma; D., Brock, and D., Engels, (2001). "Radio Frequency Identification and the Electronic Product Code", IEEE Micro, Volume: 21, Number: 6, pp. 50–54.

[13] S., Seidman; P., Ruggera; R., Brockman; B., Lewis, and M., Shein, (2007). "Electromagnetic Compatibility of Pacemakers and Implantable Cardiac Defibrillators Exposed to RFID Readers", International Journal of RFID Technology and Applications, Volume: 1, Number: 3.

[14] A., Rogers; E., Jones, and D., Oleynikov, "Radio Frequency Identification (rfid) Applied to Surgical Sponges", Surg Endosc.

[15] A., E. Lincourt; A., Harrell; J., Cristiano; C., Sechrist; K., Kercher, and B., T. Heniford, "Retained Foreign Bodies after Surgery", The Journal of Surgical Research, Volume: 138, Number: 2, pp. 170–174.

[12] M., Pelter; K., Stephens, and D., Loranger, (2007). "An Evaluation of a Numbered Surgical Sponge Product", AORN J, Volume: 85, Number: 5, pp. 931–40.

[13] C., Fabian, (2005). "Electronic Tagging of Surgical Sponges to Prevent their Accidental Retention", Surgery, Volume: 137, Number: 3, pp. 298–301.

[14] A., Macario; D., Morris, and S., Morris, (2006). "Initial Clinical Evaluation of a Handheld Device for Detecting Retained Surgical Gauze Sponges Using Radio Frequency Identification Technology", Arch Surg.

[15] "ClearCount Medical Solutions", Available: http://www.clearcount.com

WEBSITES

1. www.ieee.org

2. www.britannica.com

3. www.geister.com

*****

A SURVEY ON ROUTING IN ENERGY CONSTRAINED WIRELESS SENSOR NETWORKS EXPLOITING DATA AGGREGATION

M. Princy[38] Bessy M. Kuriakose[39]

ABSTRACT

Wireless Sensor Networks require several sensing nodes in the network to transport the sensed data to the base station so that it obtains up-to-date information on the event. A routing protocol is often triggered in order to find a path to the sink so that the nodes can report the event to the sink.

The most important task of a routing algorithm is to decide on which set of intermediate nodes is to be chosen in order to form an efficient data forwarding path between the source and the destination. The attribute of wireless sensor network environment and the characteristics of the environment where the nodes are deployed make the routing problem very challenging. The Sensor data from different nodes in a dense region may be highly correlated. Such data when routed across a wireless sensor network leads to the data redundancy at various nodes, thereby consuming a vast amount of energy.

This paper conducts a survey on the various techniques that provide an energy efficient routing for in wireless sensor networks.

KEYWORDS

Wireless, Sensors, Networks, Algorithm, Energy, Nodes, Routing, Multi hop, Correlation, Query, Maximum Lifetime etc.

I. INTRODUCTION

Wireless sensor networks are large scale networks that consist of small sensor nodes that are densely deployed in an ad hoc fashion. Each sensor node consists of one or more microcontrollers that provide processing capability, multiple types of memory, an RF transceiver, a power source, various sensors and actuators. Sensors generate a data stream by periodically measuring the physical environment around them. This data is sent to the base station which acts as a gateway that relays the data through a wireless network to the users. As there is no differentiation of application and network layers as in the internet, individual sensors in a network does the data extraction and routing. Sensor network architecture is designed specifically for an application and is not general purpose networks.

Wireless Sensor Networks are tremendously adaptable and can be deployed to support a large diversity of applications in many different conditions, whether they are poised of fixed or mobile sensor nodes. The manner these sensors are deployed is based on the temperament of the application. In environmental monitoring and surveillance applications, sensor nodes, in general are deployed in an ad hoc fashion in order to cover the particular area to be monitored.

[38] PG Scholar, Karunya University, Tamil Nadu, India, princym@karunya.edu
[39] PG Scholar, Karunya University, Tamil Nadu, India, bessymk3@gmail.com

In health care related applications, wearable wireless devices and biologically attuned sensors can be attached to or entrenched tactically within the human body to monitor very important signs of the patient under surveillance. Once positioned, sensor nodes self-organize into a self-governing wireless ad hoc network, which requires very little or no maintenance. Sensor nodes then cooperate with each other to carry out the responsibilities of the application for which they are deployed. Regardless of the inconsistency in the objectives of sensor applications, the main task of sensor nodes is to sense and collect data from an intended province, process the data, and transmit it back to the base station where the fundamental application resides. Accomplishing this task efficiently requires the advance of an energy-efficient routing protocol to set up paths between sensor nodes and the data sink. The path selection must be in such a way that the network lifetime is maximized. The distinctiveness of the environment within which sensor nodes typically operate, tied with rigorous resource and energy insufficiency, make the routing problem very challenging.

The large scale deployment of sensor nodes is less expensive and is designed with negligible complexity. Energy is a key concern in WSNs, which must achieve an extended life span while working on limited battery reserves. Multi hop packet transmission over wireless networks is a major source of power consumption. Mitigation of energy consumption can be achieved by dynamically controlling the duty phase of the wireless sensors. The energy management problem, however, becomes particularly challenging in many mission-critical sensor applications. The necessities of these applications are such that a predestined level of sensing and communication performance constriction must be maintained concurrently. Therefore, a difficulty arises as to how to devise scalable routing algorithms that can operate capably well for an extensive range of performance constraints and design requirements. The development of these protocols is fundamental to the future of WSNs.

II. EXISTING APPROACHES

1. Spin

SPIN (Sensor Protocols for Information via Negotiation) [11] is a family of adaptive protocols with the aim to overcome the shortcomings of flooding and gossiping by utilizing data negotiation and resource adaptive methods. The motives of SPIN protocol are based on two ideas: energy efficient transmission by sending meta-data and resource adaptation. This is achieved using three types of messages namely, ADV, REQ, and DATA. ADV is used to advertise the data that a node wants to send, via broadcasting this message containing meta-data (i.e., descriptor) to all nodes in the network. REQ is sent by a node that wishes to receive some data. DATA message contains the actual sensor data along with meta-data header. SPIN is a data-centric routing technique [11] where the sensor nodes send ADV message via broadcasting for the data they have and wait for REQ messages from interested sinks or nodes. SPIN uses application specific meta-data to name the sensed data. The Advantages of this technique is that it is simple, as each node needs to know only about its single-hop neighbours. It also reduces energy consumption when

compared to flooding disseminating data as quickly as possible. It can even be run on a completely configuration less network with a small start up cost to determine the nearest neighbours.

2. Minimum Energy Routing

This technique [12] mainly concentrates on Multi-hop wireless networks that typically possess two main characteristics: limited battery power and higher communication costs. Energy-aware routing protocols for such networks choose routes that minimize the total transmission power over all the nodes in the selected path. In constant power scenarios, i.e., where the transmission power is static, each link has the same cost and the minimum-hop path is preferred. But in variable power scenarios, where the transmission power can be diverse with the distance of the link, the transmission power is in general, a function of the distance between the transmitter and receiver nodes. The link cost is higher for longer hops and hence the energy-aware routing algorithms select a path with a large number of small-distance hops.

This paper [12] discusses how the formulation of the link cost fails to capture the actual energy spent in reliable packet delivery— a more precise formulation must consider the link error rates to account for the potential cost of retransmissions needed for reliable packet delivery. Therefore, the energy cost associated with an individual path should thus reflect not only the energy spent in just transmitting a single packet across the link, but rather the "total effective energy" spent in packet delivery, which includes the energy spent in potential retransmissions as well. This is discussed in terms of two different operating models: End-to-End Retransmissions (EER) and Hop-by-Hop Retransmissions (HHR). The Metrics used to explain this are normalized energy and effective throughput.

3. Data Compression Technique

This technique [13] considers a scenario of broadcast communication in a multi-hop sensor network, in which samples of a random field are collected at each node of the network, and the goal is to obtain an estimate of the entire field for all the nodes within a prescribed distortion value. The main idea is to jointly compress the data generated by different nodes as this information travels over multiple hops to eliminate correlations in the representation of the sampled field. To achieve this, the first thing is to find out how much traffic this network generates, i.e., the statistics of the sensed data. Then the network capacity to transport all that data must be checked for sufficiency. All the nodes in the network then receive information about the measurements collected by all other nodes. It is stated in [13] that even optimal vector quantization strategies cannot compress the data enough so that the network can carry it which means the network does not scale up. This scaling analysis ignores the increase in correlations of the sensor data with the increase in the density of nodes. If the data is so highly correlated that all sensors observe the same value, at it is clear that almost no exchange of information is needed for each node to have the knowledge of the local sample and of the global statistics.

This approach discussed in this paper [13] is not based on distributed coding techniques, but instead on the use of classical source codes combined with suitable routing algorithms and re-encoding of data at intermediate relay nodes. The advantage is that it prevents congestion as it uses the source coding technique. But the compression ratio will be higher as more redundant data are pooled together at each node.

4. Directed Diffusion

Directed diffusion [14] is a data-centric and application aware routing protocol for WSNs which aims at naming all data generated by sensor nodes by attribute-value pairs. Directed diffusion consists of several elements; firstly naming where task descriptors, sent out by the sink, are named by assigning attribute-value pairs.

Secondly, interests and gradients; the named task description represents an interest that contains timestamp field and several gradient fields. Each node stores the interest in its interest cache. As the interests propagate throughout the network, the gradients are set up from the source back to the sink. Thirdly, data propagation, when the source has data propagation, when the source has data for the interest, it sends out the data to the sink along the interest's gradient path.

Finally, after the sink starts receiving low rate data events, it reinforces one particular neighbour to draw down higher data rate events. Directed diffusion assists in saving sensors' energy by selecting good paths by caching and processing data in-network since each node has the ability for performing data aggregation and caching. Directed diffusion also has its limitations in implementing data aggregation which requires deployment of synchronization techniques that is not realizable in WSNs. Also, the overhead in data aggregation involves recording information.

5. Leach

LEACH (Low Energy Adaptive Clustering Hierarchy)[15] is a self-organizing and adaptive clustering-based protocol that uses randomization of cluster-heads so that the energy load among the sensor nodes in the network is evenly distributed. The two assumptions on which LEACH is based upon are, the base station is fixed and located far away from the sensors and all nodes in the network are homogeneous and energy constrained.

The main idea is to form clusters of the sensor nodes depending on the received signal strength and use local cluster heads as routers to route data to the base station. The main features of LEACH are the localized coordination and control for cluster set-up and operation, randomized rotation of the cluster "base stations" or "cluster-heads" and the corresponding clusters and the Local compression to reduce global communication. The operation is separated into fixed-length rounds, where each round starts with a setup phase followed by a steady-state phase. The duration of a round is determined beforehand. The different phases in which LEACH algorithm works are advertisement, cluster setup, schedule creation and data transmission. This algorithm cannot be applied to time-constrained application as it results in a long latency. The number of clusters is not fixed in every round and hence it cannot be applied to large sensor networks.

6. Pegasis

PEGASIS (Power-Efficient GAthering in Sensor Information Systems)[16] is a greedy chain-based energy efficient algorithm. The key characteristics of PEGASIS are the Base Station that is fixed at a far distance from the sensor nodes, the homogeneous and energy constrained sensor nodes and static

sensor nodes. PEGASIS is based on two ideas; Chaining and Data Fusion. In this technique, each node can take the turn of being a leader of the chain, where the chain can be constructed using greedy algorithms that are deployed by the sensor nodes.

PEGASIS assumes that sensor nodes have a global knowledge of the location information about all other nodes in the network. PEGASIS outperforms LEACH by eliminating the overhead of dynamic cluster formation, minimizing the sum of distances that non leader-nodes must transmit, limiting the number of transmissions and receives among all nodes, and using only one transmission to the BS per round. This technique[16] also faces the same problems that LEACH suffers from with no scalability and hence cannot be applied to sensor network where global knowledge of the network is not easy to obtain.

6. Power Aware Routing (PAR)

In-network aggregation uses routing trees to play the central role in wireless sensor networks. Two main problems faced in the case of in-network aggregation according to [1][2][3] are i) many data packets are compressed into a single packet there by hiding the detailed information. ii) The availability of the routing tree is a matter of concern and not the optimality of the tree. This technique explained in [4] categorizes aggregation into 3 levels as full aggregation, partial aggregation and No aggregation. Full aggregation happens when the data packets from the sensing nodes are transmitted as a single data packet at the parent node after the necessary compression of the correlated data has been done. Partial aggregation keeps a limit on the amount of data that each node transmits. No aggregation happens when an intermediate node has to transmit all the incoming data towards the base station. Different optimal routing tree will be obtained as a result of each of the type of aggregation. Optimal routing tree determination has the transmission costs as well as the reception costs playing a strong role. The energy that is utilized when a sensor node is idle is almost equivalent to the energy required while receiving the packets. Therefore, the transmission and reception of data as well as the idle time has a great impact on the total energy consumption.

The paper [4] shows that, for fully aggregated, non-aggregated or partially aggregated queries, determining an optimal routing tree is NP-complete. It also proposes a routing algorithm with the performance that guarantees a worst case constant factor for fully aggregated queries and an approximation algorithm that gives a near optimal performance for non aggregated and partially aggregated queries.

The model assumes a network of N sensor nodes (i=1, 2, 3... n) and an energy e_i where i=1, 2... n. Node 1 acts as the base station as it has unlimited power supply (i.e., $e_1 = \infty$). Every node generates one unit of data per unit time. It costs one unit of energy to transmit one unit of data and c_r < 1amount of energy for each sensor. Two sensors are said to communicate with each other if they are within a comfortable range r. Sensors does not dynamically adjust their power for each transmission depending on the distance to the receiver node. The radio links in the sensor networks are assumed to be symmetric links. A time synchronization model has been assumed here for computing the aggregates. In this, time is divided into epochs that are equal sized periods. During every epoch, a leaf node transmits their data to the parent. If the depth of the routing tree is d, it requires d epochs for the data from a leaf node to reach the base station. The partial state record (PSR) size varies in the different types of aggregation.

For a fully aggregated query, the PSR size is a constant irrespective of the number of sensors. For a non- aggregated query, the PSR size closest to the root is proportional to the number of total sensors. In the case of a partially aggregated query, the PSR size increases in size as it flow towards the root, but the size is maximum independent of the number of sensors. For a fully aggregated data transmission, a routing algorithm that guarantees a worst case constant factor performance. To increase the network lifetime, a node should have as few neighbours as possible that minimize the energy consumption. It is operated in two steps, the first being the conversion of the routing problem to the MDST problem where they consider a network with nodes of unequal energies. The condition specified in this paper [4] for the power constraint to be satisfied is:

$$e_i \div (1+c_r(B_i-1)) >= T$$

Where,
e_i = Varying energies of different nodes, i=1,2,…….,N,
c_r = cost of receiving one unit of data and c_r<1,
B_i = the number of neighbors of each node,
T = lifetime.

The root node has no power constraints and hence B_{root}=N. For every node i, N-B_i auxiliary nodes are taken and connected to i. This type of graph is called an augmented graph according to the author of []. A Minimum Degree Spanning Tree (MDST) is constructed for this augmented graph. The auxiliary nodes from the resulting MDST are removed leaving the tree with a degree of at most B_i for every node. The second step involves solving the MDST problem. There is a polynomial time algorithm that can produce a spanning tree with a maximum node degree of q+1, if the maximum degree of MDST is q. The problem of energy efficiency in routing is intimately connected to the type of query that the database is expected to answer. The advantages of this paper are that the algorithms proposed to suit the different types of aggregation, yields better performance. The drawback of these methods is that there is a high overhead for route reconstruction when the query type changes and this scheme is unsuitable in multitask environments where the user interests and the type of query changes frequently.

7. Hierarchical Matching Algorithm

The paper [5] introduces an algorithm called the Hierarchical matching algorithm that considers a tree that sends information from several sources to a single sink in a network.

Data Aggregation takes place when two sensors route their data to another one, in order to achieve reduction in the transmitted number of bits. If information from j sources is routed over the same link, the total information that needs to be transmitted is f(j), where f is called the aggregation function. The main requirements assumed here in this paper is that f is concave and non-decreasing and f(0)=0.

Hierarchical matching algorithm is a simple algorithm that has the above requirements satisfied. This algorithm runs in log k phases. Each phase has two steps, Matching step and Random Selection step. The matching step involves finding a perfect match in the sub graph that has a minimum cost. For example, (x_i,y_i) represents the i^{th} matched pair. The random selection

step includes choosing one out of the (x_i,y_i) matching pairs with equal probability and removing it out of the set of source nodes(S) in the network. Hence, in each step, the size of S becomes half. After log k phases, the magnitude of S becomes 1.The result would be the union of each of the log k matching and the edge connecting the single element that is remaining in S to the sink. The set of edges in the resultant graph is the aggregation tree produced by the algorithm. The authors of [5] assume that if the function f is fixed and known well in advance, the problem can be best understood. The paper aims to construct a tree that works well simultaneously with all possible aggregation functions. This is not instantaneously obvious but can achieve a simultaneous logarithmic approximation for all the aggregation functions. Since aggregation function is not known initially, simultaneous optimization is a matter of concern in this paper. The advantage of this algorithm is that it considers all the three types of aggregation models and hence, there is no need of changing the algorithm based on the type of query. It also minimizes the total energy consumption for data gathering. The drawback lies in the fact that it is dependent only on the number of nodes providing data to the node that performs aggregation and does not take into account the correlation of data among the available data.

8. Maximum Residual Energy Path Algorithm

The algorithm proposed in [6] gives knowledge that the network lifetime can be maximized by using a link metric based on the residual energy information of the sensors and the energy expenditure at each node for the transmission of information over the wireless links. The algorithm focuses on routing packets through the paths with maximum residual energy thereby balancing the energy consumption in all the paths. The residual energy can be obtained using the path length vector method, which initially determines the shortest path from each node by using the Bellman Ford algorithm. The shortest path has the largest minimum residual energy. The shortest path calculation is done by a conventional summation method that divides the link cost function which is the energy expenditure over the link by the residual energy of the node at the head of the link.

The authors of the paper[6] suggests that in order to maximize the system lifetime the residual energy information must be considered along with the information about the energy expenditure at each sensor node during the packet transmission. The new metric introduced in the paper [6] is mainly concerned with the relative residual energy which predicts the energy consumption for the future transmissions and the link capacity. The drawback of this algorithm is the loss of some important information as it is concerned with only the maximum residual energy and not the path length vector.

9. Cluster Based Maximum Lifetime Data Gathering with Aggregation (CBMLDA)

This paper [7] on cluster based heuristic algorithm tells that the key challenge faced by data gathering in sensor networks is maximizing the network lifetime in spite of the energy constraints. This algorithm is a modification to the MLDA (Maximum Lifetime Data Gathering with Aggregation) algorithm. The Maximum Lifetime Data Aggregation (MLDA) problem is that given a collection of sensors and a base station, along with their locations and the available energy of each sensor, the MLDA problem is to find a data gathering schedule with maximum lifetime, where the sensors are allowed to aggregate incoming data packets.

In this paper, the authors have discussed about finding an efficient way to collect the data from different sensors and transmitting them to the base station, when the energy available at each sensor and the location of sensors and base station are given. This is called the Maximum Data Gathering Problem. The lifetime of a sensor network can be defined as the number of rounds of transmission that can take place until the sensor responsible for the transmission is drained of all its energy. This paper aims at finding a schedule that contributes in maximizing the network lifetime. The algorithm devised for this is called a Cluster based Maximum Lifetime Data gathering using Aggregation and it requires an input of the location of n sensors and a base station and the energy that is available initially in each sensor. The n sensors are partitioned into clusters referred to as super sensors. It also forms a super sensor that consists of the base station alone.

This enables the determination of an admissible flow network for each super sensor with the base station, by using the energy of each super sensor and the distance between each of them after which a data gathering schedule is created using the LRS protocol that initially groups the sensor nodes into clusters based on their distance from the base station. The schedule and the network lifetime are then initialized to a value of 0. For each minimum residual energy sensor node, an Aggregation super tree is chosen under which an aggregation tree is initialized to contain only the base station. An aggregation tree is computed for other sensor nodes and the schedule is updated by concatenating the present schedule of the nodes with the newly formed aggregation tree of the node. The network lifetime information is updated by incrementing the current lifetime by one. Thus, the paper aims to find a schedule that helps in the maximization of the network lifetime. A schedule, by the way, is a collection of directed trees that has a root at the base station and can span the entire number of sensors. The necessary condition for a data gathering schedule to have particularly specified lifetime is that every node in the flow network must be able to push the flow to the base station.

The advantage of the technique, apart from solving MLDA problem, is that the algorithm, based on the assumption that all sensor nodes perform data aggregation with the data of any other sensors in the network, works well in the maximization of the lifetime of the system. But this makes the sensor nodes to take up the data from any sensor that comes its way. This allows the sensors to transmit undesirable data to every other node.

10. AFST

In the AFST algorithm proposed in [8], the introduction of fusion cost allows the sensor nodes to adaptively adjust its aggregation decisions along a route. If the reduction in the transmission cost due to aggregation does not become more significant than the incurred fusion cost, the algorithm in [8] will cause the data aggregation in the intermediate nodes to halt automatically. During the route construction process, AFST dynamically assigns fusion decisions to the routing decisions to the nodes by evaluating whether data aggregation is necessary at the nodes. Like the Hierarchical matching algorithm, AFST also runs in phases where each phase includes three steps. The steps are:

i) Matching Step

The matching step involves calculating the total cost of an edge which connects two nodes, that includes the transmission cost as well as the fusion costs. The minimum cost out of the costs of all edges is selected by performing the minimum cost perfect matching.

ii) Decision Step

The decision step includes making decisions based on the benefits of data aggregation regardless of which node is selected as the point of fusion. If there is no benefit in performing data fusion, the data from all intersecting nodes will be routed to the sink through the shortest path. Then these nodes that do not take part in fusion are removed from the initial set of sources.

iii) Fusion Step

The fusion step includes selecting one node from every fusion pair as the fusion point. The non fusion node transfers its data to the respective fusion node and the non fusion node is removed from the initial set of sources. The remaining set of fusion nodes form another set which would become the set of sources for the next iteration of the algorithm. The algorithm is repeated until there is only the sink node that is left. The output of the algorithm is a routing structure that has two parts: a lower part where aggregation is always performed and an upper part where there is no aggregation.

11. Minimum Energy Data Gathering Algorithm (Mega)

In the foreign coding aggregation model, data once aggregated will not be aggregated again at successive facilities. Hence, after data aggregation, the best approach for a node sending its data to the sink is to transmit the aggregated data over the minimum energy route. MEGA [9] is an algorithm that tries to minimize the cost of transmitting aggregated and raw data jointly. The resultant topology is a permutation of two tree structures, namely coding tree and SPT or may be MER tree. The coding tree is for aggregation of raw data and MER or SPT tree is for the aggregated data. This technique is formulated only to support the foreign coding aggregation model.

12. Correlation Aware Routing (CAR)

The paper [10] discusses various types of routing that takes into account various factors like the interference, energy and data correlation. The paper is about finding most correlated routing paths for taking energy efficiency as the objective function. In each iteration, each user finds the route that gives the maximum utility so that the total energy in the network is minimized. Finding optimal solutions is NP hard, so this algorithm finds a suboptimum distributed solution. The authors of the paper define a game theoretic framework and formulate the problem as a potential game which means that the algorithm will converge to Nash Equilibrium solution; however it may not be the optimal solution.

In this paper, correlated data from several nodes are aggregated into one condensed data stream in order to ease the network load. For data aggregation, we assume the step-by-step multi-hop aggregation model that is devoid of any losses. In this method, each source node aggregates its data with that of its child nodes in a step-by-step method. The paper also discusses about setting the costs for facilities for which they consider the following parameters:

a. The energy exhausted on forwarding the data from the facility through the outgoing links,
b. The opportunity for aggregation in the nodes.

The algorithm was initialized with the minimum energy routes (MER) as an alternative for the shortest path tree (SPT). It was found that starting with MER [10] gave better results at the end compared to SPT. Starting with a star topology construction, where every node sent data directly to the destination node yielded even worse results on average. This algorithm requires information to be swapped between neighbouring nodes in the network. In order to calculate the utility function of MER, each node has to know the utility function of every potential next hop. The utility function of every potential route is calculated by summing up the energy consumption of forwarding to next hop node and the utility function of next node. One of the most important drawbacks of the best response strategies like the CAR algorithm is the computational intricacy, which is O(N2) for each iteration, where N is the network size after the construction of the initial spanning tree using Dijikstra algorithm.

To tackle this issue, for large network sizes, the search for healthier routes can be constrained on k-hop neighbourhood nodes of each source node thereby reducing the computational complexity. This is based on the examination that a node is highly correlated within a definite radius of its neighbourhood. This is an anticipated postulation for sensor networks, since the correlation decreases as the distance between nodes increases; hence the correlation confined to the local area is dominant. The paper comes to a conclusion that in contrast to MEGA and MER algorithms which perform well under the foreign coding model and uncorrelated data respectively, CAR executes well under various sensor correlations and regardless of the type of network topology used. Hence, by taking into account the correlation structure, interference and multi-hop aggregation in constructing routes, significant effective energy gains over the classic approaches can be achieved.

Table-1: Comparison of Correlated Data Routing Techniques

Technique	Objective	Aggregation	Environment Suitable	Route selection
SPIN	Energy efficient communication And Resource Adaptation	No aggregation	No specifications	Broadcasting
MER	Energy efficient routing	No aggregation	Multi hop wireless networks	Routes with small distance hops
Directed Diffusion	Energy efficient routing with data aggregation	No aggregation	Single hop or multi hop network	Query on demand data model
DCT	Eliminate data correlations	Classical source coding technique	Multi hop network	Broadcast communication

LEACH	Extend network lifetime by reducing energy consumption	Multi hop data aggregation	Cluster based network	Dynamic cluster formation
PEGASIS	Energy consumption and reduce delay	Sequential data aggregation	Chain structure	Based on greedy algorithm
PAR	Minimize Energy & Maximize lifetime	Based on type of aggregation	Static Query types	Query type dependent tree
HMA	Minimize energy consumption	Logarithmic approximation	Network with any aggregation model	Aggregation tree
CBMLDA	Maximize the network lifetime	Aggregation tree is formed	Cluster based network	tree formed in clusters
AFST	Minimize total energy cost	Determined by the algorithm	Non negligible fusion cost	Full fusion sub trees and Shortest path tree
MEGA	Minimize total energy cost	Encoding nodes	Explicit side information	Coding tree and Shortest path tree
CAR	Maximize utility & minimize energy	Multihop aggregation model	Tree based network	Iterative selection

Source: Authors Analysis.

CONCLUSIONS

Designing an efficient transmission pattern in a wireless sensor network where all sensor nodes aggregate correlated data over intermediate nodes on the route to the sink, have been addressed. Analysis of the impact of data aggregation in establishing routing paths towards the sink for the energy minimization problem is also done. This research on several techniques that are related to the efficient routing of correlated data in a wireless sensor network have shown that power conservation remains a very crucial research area for the viability of wireless services in the future. As the application of WSNs to different fields become more obvious, progression in network hardware and battery technology will pave the way to realistic cost-effective implementations of these routing protocols.

REFERENCES

[1] Jae-Hwan, Chang, and Leandros, Tassiulas, (2000). "Energy Conserving Routing in Wireless ad-hoc Networks", in Proceeding of INFOCOM.

[2] N., Sadagopan, and B., Krishnamachari, (2004). "Maximizing Data Extraction in Energy-Limited Sensor Networks", in Proc. of INFOCOM.

[3] Gil, Zussman, and Adrian, Segall, (2003). "Energy Efficient Routing in ad hoc Disaster Recovery Networks", in Proceeding of INFOCOM.

[4] C., Buragohain; D., Agrawal, and S., Suri, (2005). "Power Aware Routing for Sensor Databases" Proc. IEEE INFOCOM, Miami, FLA, March.

[5] A., Goel, and D., Estrin, (2005). "Simultaneous Optimization for Concave Costs: Single Sink Aggregation or Single Source Buy-At-Bulk", Algorithmica, Volume: 43, Number: 1–2, August, pp. 5–15.

[6] Jae-Hwan, Chang, and Leandros Tassiulas, "Maximum Lifetime Routing in Wireless Sensor Networks".

[7] K., Dasgupta; K., Kalpakis, and P., Namjoshi, (2003). "An Efficient Clustering-Based Heuristic for Data Gathering and Aggregation in Sensor Networks", Proceeding IEEE Wireless Commun. and Networking Conference, New Orleans, LA, March.

[8] Hong, Luo; Yonghe, Liu, and Sajal, K. Das, (2007). "Routing Correlated Data inWireless Sensor Networks: A Survey", IEEE Network, IEEE.

[9] P., Von Rickenbach, and R., Wattenhofer, (2009). "Gathering Correlated Data in Sensor Networks", in: DIALM-POMC '04: Proceedings of the Joint Workshop on Foundations of Mobile Computing.

[10] E., Zeydan; D., Kivanc, and C., Comaniciu, (2008). "Efficient Routing for Correlated Data in Wireless Sensor Networks", in: IEEE MILCOM'08.

[11] W., R. Heinzelman; J., Kulik, and H., Balakrishnan, (1999). "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks", Proceedings ACM MobiCom '99, Seattle, WA, August, pp. 174-185.

[12] A., Misra, and S., Banerjee, (2002). "Minimum Energy Paths for Reliable Communication in Multi-hop Wireless Networks", in: MobiHoc '02: Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing, pp. 146–156.

[13] A., Scaglione, and S., Servetto, (2005). "On the Interdependence of Routing and Data Compression in Multi-Hop Sensor Networks", in: ACM Wireless Networks, Volume: 11, pp. 149–160.

[14] Chalermek, Intanagonwiwat; Govindan, Ramesh; Estrin, D. Heidemann, and John, Silva, (2003). "Fabio, Directed Diffusion for Wireless Sensor Networking", IEEE/ACM Transactions on Networking Volume: 11, Issue: 1 (February 2003), pp 2-16, ISSN:1063-6692.

[15] W., R. Heinzelman; A., Chandrakasan, and H., Balakrishnan, (2000). "Energy-efficient Communication Protocol for Wireless Microsensor Networks", in IEEE Computer Society Proceedings of the Thirty Third Hawaii International Conference on System Sciences (HICSS '00), Washington, DC, USA, January, Volume: 8, pp. 8020.

[16] S., Lindsey, and C., S. Raghavendra, (2002). "PEGASIS: Power-efficient Gathering in Sensor Information System", *Proceedings IEEE Aerospace Conference*, Volume: 3, Big Sky, MT, March, pp. 1125-1130.

*****

LOCATION PRIVACY IN WIRELESS SENSOR NETWORKS AGAINST A GLOBAL EAVESDROPPER – A SURVEY

Chinnu Mary George[40] N. Dhinakaran[41]

ABSTRACT

In this paper, we describe Location privacy. Location privacy is extremely important in WSNs. Information on location of events or on location of base stations can be of a primary concern of an adversary. Suppose the Panda-Hunter Game where a WSN is employed to monitor endangered pandas in their habitat. It is sufficient for the adversary to find out location of sensors currently monitoring the panda to successfully localize and capture the panda. Similarly, the adversary only needs to find out location of the base station to be able to mount a physical or other DoS attack on the base station and thus inactivate the whole network. In this paper, the two main categories of privacy preserving techniques for WSN have been presented, data-oriented and context-oriented. So, the different techniques against a global eavesdropper with respect to context privacy are discussed.

KEYWORDS

Wireless Sensor, Network, Data Mining, Context Privacy, Data Privacy, Location Privacy etc.

I. INTRODUCTION

A wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants and to cooperatively pass their data through the network to a main location. The more modern networks are bi-directional, enabling also to control the activity of the sensors. The number of sensor nodes in a sensor network can be several orders of magnitude higher than the nodes in an ad hoc network. Sensor nodes are densely deployed. Sensor nodes are prone to failures. The topology of a sensor network changes very frequently. Sensor nodes mainly use broadcast communication paradigm whereas most ad hoc networks are based on point-to-point communications. Sensor nodes are limited in power, computational capacities, and memory. Sensor nodes may not have global identification (ID) because of the large amount of overhead and large number of sensors. Since large numbers of sensor nodes are densely deployed, neighbour nodes may be very close to each other. Hence, multihop communication in sensor networks is expected to consume less power than the traditional single hop communication. Currently WSN are widely used in many applications such as [1]:

- Military Applications:
  - Monitoring, tracking and surveillance of borders;
  - Nuclear, biological and chemical attack detection;
  - Battle damage assessment
- Environmental applications:
  - Flood and oceans detection;
  - Forest fire detection;
  - Precision agriculture;
- Health applications:
  - Drug administration;
  - Remote monitoring of physiological data;
  - Tracking and monitoring doctors and patients inside a hospital;
- Home applications:
  - Automated meter reading;
  - Home automation;
  - Instrumented environment;
- Commercial applications:
  - Monitoring vibration for a building structure;
  - Monitoring traffic flow and road condition;
  - Vehicle tracking and detection.

II. CHALLENGES

The challenges in WSN can be roughly classified by the following criteria: cost, hardware, mobility, security, privacy.

- **Hardware**
  One major challenge in a WSN is to produce low cost and tiny sensor nodes. There are an increasing number of small companies producing WSN hardware and the commercial situation can be compared to home computing in the 1970s. Many of the nodes are still in the research and development stage, particularly their software. Also inherent to sensor network adoption is the use very low power methods for data acquisition.

- **Software**
  Energy is the scarcest resource of WSN nodes, and it determines the lifetime of WSNs. WSNs are meant to be deployed in large numbers in various environments, including remote and hostile regions, where ad-hoc communications are a key component. For this reason, algorithms and protocols need to address the following issues:
  - Lifetime maximization;
  - Robustness and fault tolerance;
  - Self-configuration;
  - Lifetime maximization: Energy/Power Consumption of the sensing device should be minimized and sensor nodes should be energy efficient since their limited energy resource determines their lifetime. To conserve power the node should shut off the radio power supply when not in use.

- **Operating System**
  Operating systems for wireless sensor network nodes are typically less complex than general-purpose operating systems. They more strongly resemble embedded systems, for two reasons. First, wireless sensor networks are typically deployed with a particular application in mind, rather than as a

[40] PG Scholar, Karunya University, Tamil Nadu, India, gchinnu2@gmail.com

[41]Assistant Professor, Karunya University, Tamil Nadu, India, kndhina@gmail.com

general platform. Second, a need for low costs and low power leads most wireless sensor nodes to have low-power microcontrollers ensuring that mechanisms such as virtual memory are either unnecessary or too expensive to implement.

- **Privacy**
WSN are capable to automatically collect data through efficient deployment of sensor devices. They offer great benefits to users but also exhibit significant potential for misuse particularly is the privacy. Privacy is defined as the limited access to a person. Privacy can be seen in location, data, temporal, etc. They are of two types:

Location Privacy
Location privacy is a critical issue in wireless sensor network. Location privacy is thus of utmost importance, especially in hostile environments. Failure to protect such information can completely subvert network applications. Location privacy measures thus need to be developed that can prevent the adversary from determining physical locations of source sensors. Providing location privacy in a sensor network is very challenging. First, the adversary can easily intercept network traffic due to the use of a broadcast medium for routing packets. He can use information like packet generation time and packet generation frequency to perform traffic analysis and infer the locations of monitored objects and data sinks. Second, sensors are usually resource constrained.

It is not feasible to apply traditional anonymous communication techniques for hiding communication between sensor nodes and destinations. Location privacy covers the protecting location privacy at the data source and protecting location privacy at the data sink. There are two types of adversaries basically this includes: a) local adversary, and b) global adversary. Local adversary is capable of monitoring limited network at a time. Global adversary is capable of monitoring the entire network at a time.

In the Panda-Hunter Game, large array of panda-detection sensor nodes have been deployed by the Save-The-Panda Organization to monitor a vast habitat for pandas. As soon as, a panda is observed, the corresponding *source* node will make observations and report data (e.g., what the panda is doing, etc.,) periodically to the *sink* via multi-hop routing techniques. The game also features a hunter in the role of the adversary, who tries to capture the panda by back tracing the routing path until it reaches the source. As a result, a privacy-cautious routing technique should prevent the hunter from locating the source, while delivering the data to the sink.

Traffic Analysis
Traffic analysis attacks help to reveal communication patterns, which allow an adversary to deduce a location of important nodes and then to compromise or to destroy them. Three classes of the traffic analysis attack are identified in WSNs: the rate monitoring attack, the time correlation attack and the content analysis attack. In the rate monitoring attack, an adversary observes nodes sending packet rate and moves closer to the node, which has the highest sending packet rate.

In the time correlation attack, an adversary monitors a correlation in sending times between a node and its neighbours.

The adversary tries to detect which node forwards the current packet and traces the path directly to a base station. In the content analysis attack, an adversary tries to obtain valuable information (e.g., a base station location) from packet headers and payloads.

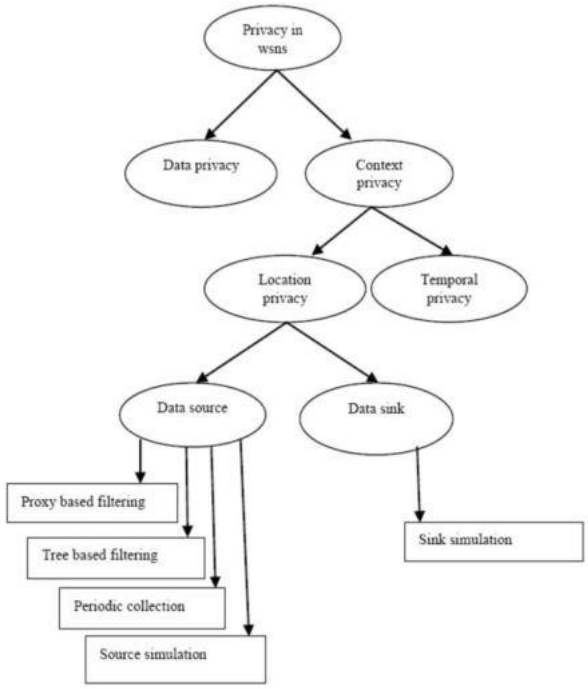
III. RELATED WORK

Privacy issue is widely explored in the field of database, networks, data mining and other field. Lots of techniques are proposed for privacy preservation such as: Cryptographic security [2], K-anonymity [3]. These techniques are use to protect data when it flows from one node to other.

A large numbers of attacks are possible in WSN such as Denial of Service attacks, The Sybil attacks, Traffic Analysis attacks, attacks against Privacy, and Physical attacks. Lot of work has done to overcome these attacks [4].

Our area of interest is on attacks against Privacy. Privacy attacks can be further classified into two broad categories data oriented attacks and context oriented attacks. Figure 1 shows the classification of privacy preservation problems in wireless sensor network. We, thus, focus on privacy preserving technique s designed to defend against a global eavesdropper.

Figure-1: Classification of Privacy Preserving Problems for WSN



Sources: Authors Compilation.

IV. PRIVACY PRESERVING TECHNIQUES

This section deals with two main branches of privacy preserving techniques data-oriented privacy and context oriented privacy respectively.

Data Oriented Privacy Protection Technique

In data-oriented privacy preserving technique main focus is on data which is collected and send to the sink. In the above mentioned example of noise and sound monitoring system; adversary can use sensor data for matching and being able to trace any person.

Context-Oriented Privacy Protection Techniques

Although privacy of data can be achieved through effective protection techniques, sensor events are so sensitive that it is needed to protect all information surrounding these sensor events. The context-oriented information includes location information which is categorized into two that is source location and sink location.

V. LOCATION PRIVACY

Location privacy plays an important role in WSN such as, location of special sensor node, data source and the base station. The famous Panda-Hunter Problem [6] in WSN sensor nodes are used to locate pandas in their local habitat. Adversary can find out location of sensor node that monitors the panda, successfully localize and capture the panda. Similarly, the location of base station is needed by adversary to mount different network attacks on the base station. Location privacy is further classified into two categories; privacy for location of data source and location of base station.

Location Privacy for Data Source

Before reviewing the existing techniques for protecting the location privacy of the data source, let us first briefly discuss the ''Panda Hunter Game", a classic formed problem, based on which (data-source) location privacy in WSN has been extensively studied in the literature [7]. In the Panda Hunter Game, a large number of panda-detecting sensors are deployed in a panda habitat. After sensors detect a panda, they will generate event messages and transmit them toward the base station. Meantime, a panda-hunter also attempts to identify the location of the data source to find the panda.

The objective of the panda hunter is to compromise the location of the data source (and thereby the panda) by analysing the traffic flow in the WSN.

Techniques

The following are the different techniques for the location privacy preservation against a global eavesdropper:

a) Proxy Based Filtering [8]

The basic idea of PFS is that a subset of sensors in the network selected as proxies collect and drop dummy messages before they reach the base station so that the problem of high communication cost of periodic packet transmission is mitigated. If one of the incoming packets to a proxy corresponds to a real event, proxy's outgoing packet carries that information; otherwise the outgoing packet is a dummy one. In another view, a proxy acts like a base station, it collects packets from other sensors but since it does not have a direct connection to the outside world, the information it collects should be sent to the base station again in a periodic fashion to preserve privacy properties. Hence the data is hidden in a controlled fashion to an outsider so that any useful information cannot be extracted from the traffic patterns observed in the network data flow (i.e., a global eavesdropper cannot determine whether an event-triggered activity is happening in the network and cannot single out any particular node as the source node).

Tree Based Filtering [8]

TFS, the second scheme, allows filtering at multiple proxies. In TFS, proxies form a tree rooted at the base station with each proxy having a parent node and possibly multiple child nodes. Parent nodes aggregate traffic originated by child nodes and child leaf nodes aggregate data coming from ordinary sensors.

Periodic Collection [9]

In periodic collection method has every sensor node independently and periodically send packets at a reasonable frequency regardless of whether there is real data to send or not. To enable this, each sensor node has a timer which triggers an event every seconds as well as a First-in-First-out (FIFO) queue of size q for buffering received packets that carry real data reports. When an event is triggered by the timer, the node checks if it has any packet in its queue. If so, it dequeues the first packet, encrypts it with the pair wise key it shares with the next hop, and forwards it to that next node. Otherwise, it sends a dummy packet, with a random payload, that will not correctly authenticate at the next hop. Since every sensor node only accepts the packets that correctly authenticate, dummy packets do not enter the receiver's queue. When the queue at a sensor node is full, it will stop accepting new packets.

Source Simulation [9]

Each of them will generate a traffic pattern similar to that of a real object. Both the adversary and the defender have a model of this random movement pattern. After network deployment, each virtual object is treated like a real object, as sensors detect it and send the object's information to the destination.

The protocol works in rounds. In every round, the node simulating the fake object will randomly pick a sensor node in its neighbourhood (including itself) and ask this node to simulate the real object in the next round. In this way, there will be multiple movement patterns similar to that of real objects.

Sink Simulation [9]

In the destination simulation approach, fake destinations will be simulated in the field. Each of them will receive traffic similar to the traffic received by a real destination. To achieve this, we make no distinction between fake and real destinations when sensors send data.

During deployment, we place real destinations and select locations where fake destinations are to be simulated. A few sensors will be used as the fake destinations. It is also required that each real destination have a fake destination simulated in its communication range.

Tables-1: Comparison between Different Techniques

Techniques	Location	Privacy Preservation Efficiency	Adversary	Message Overhead	Power Consumption	Accuracy	Delay
Proxy Based Filtering	Sink	Good	Global	No Extra Over Head	High	No Influence Data Arrival Accuracy	Extra Head
Tree Based Filtering	Sink	Good	Global	Small	Medium	No Guaranteed Data Arrival	Delay Can Increase
Periodic Collection	Sink	Optimal	Global	High	High	No Guaranteed Data Arrival	Delay Can Increase
Source Simulation	Sink	Excellant	Global	Fair	Low	No Guaranteed Data Arrival	Delay Increase While Improving The Safety Period

Sources: Authors Compilation.

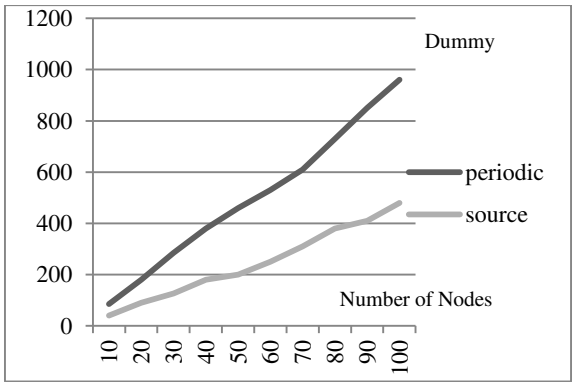
VI. COMPARISONS

In this we compare all the privacy preserving techniques that have been reviewed in this paper. Table-1 depicts the performance of privacy preservation techniques in WSNs. We evaluate their performances in metrics: privacy, accuracy, delay time, and power consumption, location, adversary, message overhead, scalability. Privacy refers to the degree of privacy protection provided by the reviewed techniques. The accuracy measure covers two perspectives: (i) the accuracy of the data obtained by the base station; and (ii) the availability of the (intended) data to the base station (i.e., whether the data can be delivered to the base station). The delay time includes both the computation and communication time of data transmission at the intermediate sensors. The power consumption measure focuses on the additional messages required for transmission (i.e., additional energy consumed) in the WSN. There are some open issues for future research. As the network behavior changes so sensor nodes should be designed accordingly, along with it there can be improvement in every technique with respect to overhead, power, accuracy.

VII. SIMULATION EVALUATION

The Graph-1 shows that the message overhead in periodic collection is high as compared to source simulation. The periodic simulation can applied to applications that collect data at low rate and no strict requirements on data delivery latency. Source Simulation can be applied to real-time applications.

Graph-1: Message Overhead of Periodic Simulation & Source Simulation



Sources: Data Analysis.

VII. CONCLUSIONS

This paper presents a review of privacy-preserving techniques for wireless sensor networks (WSN) against a global Eavesdropper. Two main categories of privacy preserving techniques have been presented; data oriented and context-oriented respectively. The existing techniques have been compared in terms of privacy preservation efficiency, overhead, delay, power consumption adversary, location.

In regard to the future scope we assume that a global eavesdropper cannot comprise the sensor nodes. However in practice they have this ability to comprise a subset of the nodes and then perform traffic analysis and get knowledge from that of the insiders.

VIII. REFERENCES

[1] K., Sohraby; D., Minoli, and T., Znati, (2007). "Wireless Sensor Network: Technology, Protocols & Applications", John Wiley & Sons, pp. 10-11.

[2] C., Ozturk; Y., Zhang, and W., Trappe, (2004). "Source-Location Privacy in Energy Constrained Sensor Network Routing", In Proceedings of the 2nd ACM workshop on Security of Adhoc and Sensor Networks.

[3] R., Agrawall; A., Evfimievski, and R., Srikant, (2003). "Information Sharing Across Private Databases", Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data, pp. 86–97.

[4] L., Sweeney, (2002). "K-anonymity: A Model for Protecting Privacy, International Journal on Uncertainty, Fuzziness and Knowledge based Systems, 2(2), 557–570, pp. 86– 97.

[5] Na, Li; Nan, Zhang; Sajal, K. Das, and Bhavani, Thuraisingham, (2009). "Privacy Preservation in Wireless Sensor Networks: A State-of-the-art Survey", Ad Hoc Networks, 7, pp. 1501–1514.

[6] Jean, Franois Raymond, (2001). "Traffic Analysis: Protocols, Attacks, Design Issues and Open Problems", In Proceedings of International Workshop on Design Issues in Anonymity and Unobservability, pp. 10-29, Springer-Verlag New York, Inc.

[7] Celal, Ozturk; Yanyong, Zhang, and Wade, Trappe, (2004). "Source Location Privacy in Energy-Constrained Sensor Network Routing", SASN '04: Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks, pp. 88-93, New York, NY, USA, ACM.

[8] Y., Xi; L., Schwiebert, and W., S. Shi, (2006). "Preserving Source Location Privacy in Monitoring - Based Wireless Sensor Networks", Proceedings of the 20th International Parallel and Distributed Processing Symposium (IPDPS 2006), April.

[9] K., Mehta; Donggang, Liu, and M., Wright, (2007). "Location Privacy in Sensor Networks against a Global Eavesdropper", IEEE International Conference on Network Protocols, ICNP 2007, October, pp. 31-323.

[10] Jing, Deng; Richard, Han, and Shivakant, Mishra (2004). "Intrusion Tolerance and Anti-Traffic Analysis Strategies for Wireless Sensor Networks", DSN '04: Proceedings of the 2004 International Conference on Dependable Systems and Networks, pp. 637-646, Washington, DC, USA, IEEE Computer Society.

[11] Yi, Ouyang; Zhengyi, Le; Guanling, Chen; James, Ford, and Fillia, Makedon, (2006). "Entrapping Adversaries for Source Protection in Sensor Networks", WOWMOM '06: Proceedings of the 2006 International Symposium on World of Wireless, Mobile and Multimedia Networks, pp. 23-34, Washington, DC, USA, IEEE Computer Society.

[12] Jing, Deng; Richard, Han, and Shivakant, Mishra, (2004). Intrusion Tolerance and Anti-Traffic Analysis Strategies for Wireless Sensor Networks", DSN '04: Proceedings of the 2004 International Conference on Dependable Systems and Networks, pp. 637-646, Washington, DC, USA, IEEE Computer Society.

[13] Y., Jian; S., Chen; Z., Zhang, and L., Zhang, (2007). "Protecting Receiver-Location Privacy in Wireless Sensor Networks", May, pp. 1955-1963.

*****

PEZZOTTAITE JOURNALS MESSAGE TO AUTHORS

We require that, prior to publication; authors make warranties to these effects when signing their Agreements.

An author must not submit a manuscript to more than one journal simultaneously, nor should an author submit previously published work, nor work which is based in substance on previously published work.

An author should present an accurate account of research performed and an objective discussion of its significance, and present sufficient detail and reference to public sources of information so to permit the author's peers to repeat the work.

An author must cite all relevant publications. Information obtained privately, as in conversation, correspondence, or discussion with third parties, should not be used or reported in the author's work unless fully cited, and with the permission of that third party.

An author must make available all requisite formal and documented ethical approval from an appropriate research ethics committee using humans or human tissue, including evidence of anonymisation and informed consent from the client (s) or patient (s) studied.

An author must follow national and international procedures that govern the ethics of work done on animals.

An author must avoid making defamatory statements in submitted articles which could be construed as impugning any person's reputation, for example, making allegations of dishonesty or sharp practice, plagiarism, or misrepresentation; or in any way attacking a person's integrity or competence.

An author must ensure all named co-authors consent to publication and being named as a co-author, and, equally, that all those persons who have made significant scientific or literary contributions to the work reported are named as co-authors.

Additionally, the author understands that co-authors are bound by these same principles.

(sd/-)
(Editor-In-Chief)

FOR ANY CLARIFICATION OR SUGGESTION, WRITE US:

Editor-In-Chief
Pezzottaite Journals,
24, Saraswati Lane, Bohri, Near Modern Dewan Beverages,
Jammu Tawi – 180002, Jammu and Kashmir, India.
(Mobile): +91-09419216270 – 71
editorinchief@pezzottaitejournals.net
contactus@pezzottaitejournals.net

SURFACE HARDNESS & TENSILE BEHAVIOUR OF Ni-Cr-Mo ALLOYS

V. K. Murugan[42] P. Koshy Mathews[43]

ABSTRACT

Low carbon steel is easily available and cheap having good material properties that are acceptable for many applications. The mechanical properties such as ductility, toughness, strength, hardness and tensile strength can easily be modified by heat treating the low carbon steel to suit a particular design purpose. The specimen of steel was purchased from market and checked for the selected mechanical behavior of Ni-Cr-Mo alloy steel when it is before and after the heat treatment processes. The steel samples were heat treated in electric furnaces (for carburizing, hardening and tempering processes) at different temperature levels and holding time and then cooled in different media. The specimens are heat treated at 900^0C for 1hr and the specimen then undergone to tempering process for different temperature level such as 250^0C, 350^0C, 450^0C and 550^0C for different tempering time (60min, 75 min., and 90 min). The mechanical behavior of the samples was investigated using universal tensile testing machine for tensile test and Brinell hardness method for hardness testing. In this study, the aim is to evaluate the effect of heat treatment processes on the hardness and tensile behavior of low carbon alloy steels. Results showed that the mechanical properties of low carbon alloy steel can be changed and improved by various heat treatments for a particular application.

KEYWORDS

Carbon, Steel, Alloy, Tenstile, Hardness etc.

I. INTRODUCTION

Steel is an alloy of iron with definite percentage of carbon ranges from 0.15-1.5%, plain carbon steels are those containing [1].Plain carbon steels are changed on the basis of their carbon content as their major alloying element in carbon. Steels with carbon content varying from 0.25% to 0.65% are classified as medium carbon, while those with carbon content less than 0.25% are termed low carbon. The carbon content of high carbon steels usually ranges within 0.65 – 1.5%. Steel with low carbon content has the same properties as iron, soft but easily formed. As carbon content rises, the metal becomes harder [2] and stronger but less ductile and more difficult to weld. The steel grades are usually required for the production of body parts of vehicles, trucks, and machineries [3]. Such steels offer impressive mechanical properties and the advantage of reduced cost, excellent formability and surface finish among other HSLA Steels [4]. The steel grades commercially utilized for these applications are designed with compositions usually containing micro-alloying elements like molybdenum, niobium, nickel, and chromium among others, which function as strengtheners, phase stabilizers, formability enhancers [5]. Steel which consist of ferrite and martensite / bainite [6] and have shown superior combinations of strength, toughness and plasticity over conventional high strength low alloy steels [7]. The heat treating is an easy way to improve metals and make them more versatile. The conversion of primary products in to secondary finished or semi-finished components can take place by one of several alternative routes. The technique used to obtain the secondary shape, it is almost always heat treated and finish machined .Processing Operations can be classified in to the following four categories: a) Basic Processes b) Secondary Processes c) Operations to enhance physical properties, d) Finishing operations. In the operations to enhance physical properties do not perceptibly change the physical geometry of the work part. The physical properties of the material are improved with the help of heat- treatment processes. Heat treatment is also used to increase the strength of materials by altering some certain manufacturability objectives especially after the materials might have undergo major stresses like forging and welding. Also the heat treatment develops hardness, softness, and improves the mechanical properties (such as tensile strength, yield strength, ductility, corrosion resistance and creep rupture. These processes are also help to improve [8] machining effect, to improve metals and make them to versatile. Heat treatment operation is a means of controlled heating and cooling of materials in order to effect changes [9] in their mechanical properties. The heat treatment processes is to modify the microstructure and consequently change the properties (10) of the work piece throughout. Heat treatment involves the application of heat, to a material to obtain desired material properties [11]. During the heat treatment process, the material usually undergoes phase micro structural and cryptographic changes. Heat treatment is a combination of timed heating and cooling applied to a particular metal or alloy in the solid state in such ways as to produce certain microstructure and desired mechanical properties (hardness, toughness, yield strength, ultimate tensile strength, young's modulus ,percentage of elongation and percentage of reduction).Annealing ,normalizing, hardening and tempering are the most important heat treatments often used to modify the microstructure [12]and mechanical properties of engineering materials steels. Heat Treatment consists of processes like Annealing, Normalizing, hardening, Tempering, etc. The demand in the market arises from defense, railways, automobile, fasteners, machine tool, fabricating parts in manufacturing companies. The automobile and many other industries in the manufacturing sector are constantly in demand of steel possessing a good compromise of toughness and plasticity at high strength levels [13]. Hardened sample had the highest tensile strength and hardness with lowest ductility and impact strength when compared to other heat treated samples. Hardening is strongly recommended when the strength and hardness [14] are the prime desired properties in design. The increase in tensile strength and yield strength with increase in the volume percent martensite is consistent with observations of Sun and Pugh [15] and Kumar et al [16]. Also the tensile properties and fatigue resistance [17] can be effectively enhanced for the carbon steel grades by the adoption of this processing heat-treatment condition which makes it suitable for high strength and structural application. The Steel developed by quenching followed by tempering process at a desired temperature [18] has the highest ultimate tensile strength with excellent combination of impact strength, ductility and hardness which is very attractive for structural use. The ultimate tensile strengths are steadily decrease by increasing tempering time and temperature. And also the steel developed by quenching with tempering process has the highest ultimate

[42]Associate Professor, Kalaivani College of Technology, Tamil Nadu, India, 27murugan1966@gmail.com
[43]Dean (R&D), Kalaivani College of Technology, Tamil Nadu, India, pkoshymathews@yahoo.co.in

tensile strength [19] and yield ratio with excellent combination of impact strength, ductility and hardness which is very attractive for structural use.

The ductility of the samples is measured by the tensile test. The percent of elongation is upward trend in the increment of tempering time and [20] temperature. This brittleness is therefore removed by tempering. Tempering results in a desired combination of hardness, ductility, toughness, strength and structural stability. The strength of the steel may be improved by quenching followed by tempering [21] with some compromise on toughness.

II. EXPERIMENTAL PROCEDURE

To evaluate the tensile behavior of 20Ni55Cr50Mo20 alloys on the low carbon steel, the investigation was carried out thus; (i) Chemical composition and Preparation of the specimens (ii) Heat treating the low carbon steel specimens, (iii) Determination of mechanical properties (hardness, tensile strength) of the low carbon steel specimens (iv) Finally analyze its mechanical behavior of low carbon steel specimens after the heat treatment completed.

2.1 Chemical Composition and Preparation of Specimens

The first and foremost job for the experiment is the specimen preparation. The material used for this study is a low carbon steel with carbon content of 0.18 % carbon. Sample of 20Ni55Cr50Mo20 steel bar with 25mm diameter and 1m length was purchased. The chemical composition of the sample is noted in the following Table-1

Table-1: Indicates Chemical Composition of 20Ni55Cr50Mo20 Steel.

C%	Si%	Mn%	P%	S%	Cr%	Mo%	Ni%
0.18	0.15	0.70	0.035	0.040	0.50	0.20	0.55

Sources: Secondary Data.

The material used for this study is a low carbon steel and the specimens were then prepared for a tensile test using a standard format of ASTM. Tensile tests were carried out by a universal testing machine using using JIS No.14A type specimens.

2.2 Heat treatment processes

After preparation of the specimen samples from the low carbon steel, it was taken to the furnace for the heat treatment operations. To commence the operation, the furnace was initially calibrated to determine the furnace operating temperature based on the pre-set furnace temperature. Representative samples of low carbon steel were subjected to heat treatment processes. The various forms of the heat treatment processes were stated below.

2.2.1 Hardening Process

This material is low carbon steel, hence the specimens are sent to carburizing process carburizing at 8800C) in a depth of 2mm case-depth. After the carburizing process the specimen utilized with hardening process. Hardening is carried out by quenching steel, which is cooling it rapidly from a temperature above the transformation temperature. The specimens to be hardened were placed inside the furnace and heated to a temperature of 9000C.At this temperature; there is transformation of the steel to austenite. The total heating time should be just enough to attain uniform temperature through the section of the part to enable not only the completion of phase transformation, but also to obtain homogeneous austenite.

2.2.2 Tempering Process

After the hardening of the hardened specimen is follow by the tempering. The tempering of the quenched specimens was also carried out in a muffle furnace for 1 hour. The selective steel is tempered in the temperature range of 250^0C to 550^0C in four stages. The first to fourth stages tempering involves under the temperature of 250^0C, 350^0C, 450^0C, 550^0C for tempering time of 60, 75 and 90 minutes respectively.

III. DETERMINATION OF MECHANICAL PROPERTIES

The mechanical property depends largely upon the various forms of heat treatment operations and cooling rate [22]. Mechanical Properties of the treated and untreated samples were determined using standard methods. After the specimens had been heat treated, the specimens were loaded into a 2000-Kg Auto instrument Universal Testing Machine. The tensile test is widely used for measuring the stiffness, strength and ductility of a material [23]. The tensile test were carried out on them to determine the mechanical properties of the steel and compare it with the non heat treated specimen which was also subjected to the same tensile test. Yield strength, tensile strength, and % of elongation were determined. The various heat treated samples were taken for the tensile test in the universal testing machine. The tensile test was carried out on them to determine the mechanical properties of the steel.

3.1 Hardness Testing

The hardness test is used as quick, in expensive method of assessing the mechanical properties of a material. The hardness of as received and heat-treated specimens were measured with the aid of Brinell hardness testing machine. The hardness of a material is determined by pressing an indenter into its surface and measuring the size of the impression. This machine measures the resistance to penetration by measuring the depth of impression and the hardness is indicated directly on the scale attached to the machine.

3.2 Tensile Strength Testing

After the successful heat treatment operation, the various heat treated samples were taken for the tensile test. The tensile tests were conducted by using an Auto instrument Universal Testing Machine. Each of the specimens was inserted one after the other into the machine jaws and having fastened the specimen properly at both ends, tensile test up to the fracture limit was carried out. The heat treated specimens were checked and to obtain the elongation, ultimate tensile strength, yield strength.

IV. RESULTS AND DISCUSSION

The mechanical Properties of the steel specimens after various heat treatment processes are shown in following tables. The chemical composition of test sample is C%-0.18, Si%-0.15, Mn%-0.70, P%-0.035, S%-0.040, Cr%-0.50, Mo%-0.20, Ni%-0.55. Its equivalent grade agrees with AISI 8620(DIN) Standard Specification [24]. The heat treated specimens were now subjected to hardness test, using standard brinell hardness testing machine, which is calibrated in unit of BHN and tensile

strength were found with the help of universal testing machine. The resulting values are obtained from the hardness test and tensile test were plotted the curves/Figures 1 to 5 showing the range of tempering temperature and the variation of hardness value, tensile strength, yield strength and elongation.

The data generated from these graphs for each of the heat treated specimens processes output result in mechanical properties were analyzed. The specimen hardness properties obtained were shown in Table-2.

Table-2: Hardness for Heat Treated Specimens under 250^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Hardness in BHN
1	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	60	461
2	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	75	429
3	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	90	375

Sources: Computed Data.

Table-3: Tensile Properties for Heat Treated Specimens under 250^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Tensile Strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	60	548
2	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	75	543
3	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	90	412

Sources: Computed Data.

Table-4: Yield strength for Heat Treated Specimens under 250^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Yield strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	60	334
2	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	75	331
3	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	90	288

Sources: Computed Data.

Table-5: Elongation for Heat Treated Specimens under 250^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Elongation (In %)
1	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	60	9.7
2	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	75	12.3
3	Hardening ^{0}C at 900^{0}C & Tempered at 250^{0}C	90	22.9

Sources: Computed Data.

Table-6: Hardness for Heat Treated Specimens under 350^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Hardness in BHN
1	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	60	409
2	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	75	363
3	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	90	341

Sources: Computed Data.

Table-7: Tensile Properties for Heat Treated Specimens under 350^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Tensile Strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	60	523
2	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	75	428
3	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	90	397

Sources: Computed Data.

Table-8: Yield strength for Heat Treated Specimens under 350^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Yield strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	60	315
2	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	75	307
3	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	90	271

Sources: Computed Data.

Table-9: Elongation for Heat Treated Specimens under 350^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Elongation (In %)
1	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	60	12.1
2	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	75	15.4
3	Hardening ^{0}C at 900^{0}C & Tempered at 350^{0}C	90	25.2

Sources: Computed Data.

Table-10: Hardness for Heat Treated Specimens under 450^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Hardness in BHN
1	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	60	352
2	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	75	311
3	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	90	277

Sources: Computed Data.

Table-11: Tensile Properties for Heat Treated Specimens under 450^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Tensile Strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	60	497
2	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	75	428
3	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	90	382

Sources: Computed Data.

Table-12: Yield strength for Heat Treated Specimens under 450^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Yield strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	60	297
2	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	75	307
3	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	90	255

Sources: Computed Data.

Table-13: Elongation for Heat Treated Specimens under 450^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Elongation (In %)
1	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	60	14.4
2	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	75	15.4
3	Hardening ^{0}C at 900^{0}C & Tempered at 450^{0}C	90	27.5

Sources: Computed Data.

Table-14: Hardness for Heat Treated Specimens under 550^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Hardness in BHN
1	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	60	293
2	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	75	262
3	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	90	241

Sources: Computed Data.

Table-11: Tensile Properties for Heat Treated Specimens under 550^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Tensile Strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	60	348
2	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	75	198
3	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	90	251

Sources: Computed Data.

Table-12: Yield strength for Heat Treated Specimens under 550^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Yield strength (N/mm^2)
1	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	60	234
2	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	75	238
3	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	90	198

Sources: Computed Data.

Table-13: Elongation for Heat Treated Specimens under 450^{0}C Tempering Temperature

Sl. No.	Process Specification	Tempering Time in min	Elongation (In %)
1	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	60	20.5
2	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	75	24.9
3	Hardening ^{0}C at 900^{0}C & Tempered at 550^{0}C	90	27.7

Source: Computed Data.

V. CONCLUSIONS

From the various experiments carried out, the following observations were made. More is the tempering temperature; less in the hardness, more in the softness (ductility) induced in the quenched specimen. More is the tempering time (keeping the tempering temperature content), more is the ductility induced in the specimen.

This clearly implies that UTS and also to same extent the yield strength decreases with increase in tempering time where as the ductility (% elongation) increases. Based on the given tempering time, as increase in the tempering temperature decreases the UTS value and the yield strength of the specimen where as on the often hand increasing the % elongation and hence the ductility.

Ductility is the only the criteria of tempering at high temperature for 90 minutes gives the best result among all tempering experiments.

From the various result obtained and concluded that the mechanical properties vary depending upon the various heat treatment processes. Hence depending upon properties and application required we should go for a suitable heat treatment process.

VI. REFERENCES

[1] John, V. B., (1980), Introduction to Engineering Materials, 2nd Edition, Macmillan Publishing Company Ltd., pp. 321-324.

[2] Smith, W. F., and Hashemi, J., (2006). Foundations of Materials Science and Engineering, 4th Edition, McGraw-Hill, pp. 28-36.

[3] B., Gajda; and A., K. Lis, "Intercritical Annealing with Isothermal Holding of Trip CMnAlSi Steel", Journal of

Achievements in Materials and Manufacturing, Volume: 20, Issues: 1-2, pp. 439-442.

[4] U., Liedi, and E., A., (2002). Traint: Comp. Mater.Sci.25 122.DOI:10.1016/S0927-0256 (02)00256-2.

[5] I., Nobuyuki; S., Nobuv, and K., Joe, (2008). "Development of ultra-High Strength Line Pipes with Dual-phase Microstructure for High Strain Application", JFE Technical Report, Number: 12, pp.15-19.

[6] A., Bayram; A., Uguz, and M., Ula, (1999). "Effects of Mechanical Properties of Dual-phase Steels", Materials Characterization, 43, pp. 259-269.

[7] Y, Son; Y., K. Lee; K., T. Park; C., S. Lee, and D., H. Shin, (2005). "Ultra-fine Grained Ferrite-Martensite Dual Phase Steels Fabricated via Equal Channel Angular Pressing: Microstructure and Tensile Properties', Acta Materialia, Volume: 53, Issue; 11, pp. 3125-3134.

[8] Min, Shan HTUN; Si, thu KYAW, and Kay, thi LWIN, (2008). "Effect of Heat Treatment on Microstructures and Mechanical Properties of Spring Steel", Journal of Metals, Materials and Minerals, Volume: 18, Number: 2, pp.191-1978.

[9] Adnan, Calik, (2009). "Effect of cooling rate on Hardness and Micro structure of AISI 1020, AISI 1040 and AISI 1060 Steels", Int .J. of Physics Sciences, Volume: 4(9), pp. 514-518.

[10] Ullah., A.; Ikram, N., and Salam, A., (2008). "Improvement in Tensile Properties of a Structural Steel by Heat Treatment", International Journal of Material Science, Volume: 3, Number: 1, pp.89-94.

[11] Rajan, T. V.; Sharma, C. P., and Sharma, A., (1989). Heat Treatment Principles and Techniques, New Delhi: Prentice Hall of India Private Limited, pp. 36-58.

[12] Raymond, A., and Higgins B., (1985), Properties of Engineering Materials, Hooder and Stonghton.

[13] M., Yano; S., Kitamuna; T., Iharashima; T., Inamoto; K., Azuma, and H., Nagahama, (1999). "Recent Advances in Ultra Low-Carbon Steel Refining Technology by Vacuum Degassing Processes", Nippon Steel Technical Report, Number: 61, VDC 669, pp. 18-982.

[14] Fadare, D. A.; Fadara, T. G., and Akanbi, O. Y., (2011). "Effect of Heat Treatment on Mechanical and Micro structure of NST 37-2 Steel", Journal of Minerals & Materials Characterization & Engineering, Volume: 10, Number: 3, pp. 209-308.

[15] S., Sun, and M., Pugh, (2002). Properties of Thermo Mechanically Processed Dual-Phase Steels Containing Fibrous Martensite", Materials Science and Engineering A, Volume: 335, Issue: 1-2, pp. 298-308.

[16] A., Kumar; S., B., Singth, and K., K. Ray, (2008). "Influence of Bainite/Martensite-Content on Tensile Properties of Low Carbon Dual Phase Steels", Materials Science and Engineering A, 474, pp. 270-282.

[17] K., K., Alaneme; S., Ranganathan, and T., Mojisola, (2010). "Mechanical Behaviour of Duplex Phase Structures in A Medium Carbon Low Alloy Steel", Journal of Minerals & Materials Characterization & Engineering, Volume; 9, Number: 7, pp. 621-633.

[18] Daramola, O. O.; Adewuyi, B. O., and Oladele, I. O., (2010). "Effects of Heat Treatment on the Mechanical Properties of Rolled Medium Carbon Steel", Journal of Minerals & Materials Characterization & Engineering, Volume: 9, Number: 8, pp.693-708.

[19] Daramola, O. O.; Adewuyi, B. O., and Oladele, I. O., (2010). "Effects of Heat Treatment on Mechanical Properties of Rolled Medium Carbon Steel", Journal of Minerals & Materials Characterization & Engineering, Volume: 9, Number: 8, pp. 693-708.

[20] Rao, P. N., (2011). "Manufacturing Technology", Volume: 1, New Delhi: McGraw Hill, pp.11-53.

[21] Lakhtin., Y., (1998). "Engineering Physical Metallurgy", Mascow: MIR Publishers, pp.14-55

[22] Senthilkumar., T., and Ajiboye, T. K., (2012). "Effect of Heat Treatment Processes on the Mechanical Properties of Medium Carbon Steel", Journal of Minerals & Materials Characterization & Engineering, Volume: 11, Number: 2, pp.142-152.

[23] Martin., J. W, (2009). "Materials for Engineering", Cambridge: Wood Head Publishing Limited, England, pp. 38-42

[24] Prabhudev, K. H., (1988). Hand book of Heat Treatment of Steels. Bangalore: Tata McGraw-Hill Publishing.

*****

ISOLATED TAMIL SPEECH RECOGNITION SYSTEM BASED ON CMU SPHINX

Janani J.[44] Mohanapriya N. [45]

ABSTRACT

Speech is one of the most complex signals and the powerful tool for communication. Speech recognition has been a long desire of the scientists that the machine should recognize the speech of the human beings either for the machine to function on voice commands or for giving a text output of the speech. In this paper we present the creation of Tamil version of Automated Speech Recognition System (ASR).This system is based on the open source Sphinx-4, from the Carnegie Mellon University, which is a speech recognition system based on discrete Hidden Markov Models (HMMs). We investigate the changes that must be made to the model to adapt Tamil speech recognition.

KEYWORDS

Signals, ASR, Hidden Markov Models (HMMs), Speech Recognition etc.

I. INTRODUCTION

Automatic Speech Recognition (ASR) is a technology which has a wide area of applications: Command recognition (Voice user interface with the computer), Dictation, Interactive Voice Response, it can be used to learn a language. ASR can help also, handicapped people to interact with society. It's a technology which makes life easier and has much to offer for the future [1]. Speech is the human's most efficient method of communication.

The pronunciation of words and the style of speaking vary from person to person and place to place. It is thus a complex process to recognize speech with person independent and place independent. Further the environment where speech occurs is normally polluted with noise. Hence natural speech to text recognizing process will have recorded speech signal with noises. The recognizing process need to have a noise filtering mechanism. View the importance of ASR too many systems are developed, the most popular are: Dragon Naturally Speaking, IBM Via voice, Microsoft SAPI. Open source speech recognition systems are available too, such as HTK [2], ISIP [3], AVCSR [4] and CMU Sphinx-4 [5-7]. We are interested to this last, which is a Hidden Markov Models (HMMs) based system [8].

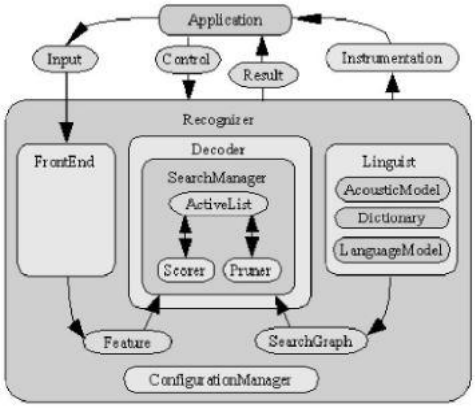
II. SYSTEM PRESENTATION

Sphinx-4 speech recognition system has been jointly developed by the Carnegie Mellon University, Sun Rosems Laboratories and It subish EltResearch Laboratories (M.E.R.L). It has been built entirely in Java TM programming language. Since the born the Sphinx Group is dedicated to release a set of reasonably mature, speech components. Those provide a basic technology level to anyone interested in creating speech recognition systems. Since 2000, first with CMU Sphinx I, CMU Sphinx II, CMU SphinxTrain then CMU Sphinx III and CMUSphinx-4, a large part of CMU Sphinx project has been made available as open source packages [9-11].SphinxTrain is the acoustic training environment for CMU Sphinx (for sphinx2, sphinx3 and sphinx4), which was first publicly released on June 7th, 2001. It is a suite of programs, script and documentation for building acoustic models from data for the Sphinx suite of recognition engines. With this contribution, people should be able to build models for any language and condition for which there is enough acoustic data.One cannot proceed to the recognition without having an acoustic model, which is necessary to compare the data coming from FrontEnd (see Figure-1).This model should be prepared using Sphinx Train tool.

2.1 Architecture

The Sphinx-4 architecture has been designed with a high degree of flexibility and modularity. Each labelled element in Figure-1 represents a module that can be easily replaced, allowing researchers to experiment with different module implementations without needing to modify other portions of the system. The main blocks in Sphinx-4 architecture are frontend, decoder and Linguist.

Figure-1: Sphinx-4 Architecture: Main blocks are Front End, Decoder, and Linguist.



Sources: Authors Compilation.

FrontEnd: Parameterize an input signal (e.g. audio) into a sequence of output features. It performs digital signal processing (DSP) on the incoming data.

Feature: The output of the front end is feature, which are used for decoding in the rest of the system.

Linguist: Or knowledge base, it provides the information the decoder needs to do its job. Is made up of three modules which are:

- **Acoustic Model:** Contains a representation (often statistical) of a sound, often created by training using lots of acoustic data.

[44] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, jananij.cse@gmail.com
[45] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, nmohanapriyabe@gmail.com

- **Dictionary:** Responsible for determining how a word is pronounced.

- **Language Model:** Contains a representation (often statistical) of the probability of occurrence of words.

Search Graph: The graph structure produced by the linguist according to certain criteria (e.g.the grammar), using knowledge from the dictionary, the acoustic model, and the language model.

Decoder: The decoder is the main block in the sphinx-4 system, which performs the bulk of the work. It reads features from the front end, couples this with data from the knowledge base and feedback from the application, and performs a search to determine the most likely sequences of words that could be represented by a series of features.

2.2 Installation

2.2.1 Sphinx-4

Sphinx-4 can be downloaded either in binary format or in source codes [12]. It was compiled and tested on several versions of Linux and on Windows operating systems. Running, building and testing Sphinx-4 requires additional software:

- Java 2 SDK, Standard Edition 5.0 [13].
- Java Runtime Environment (JRE).
- Ant: the tool to facilitate compilation and the implementation of sphinx-4 system [14].

2.2.2 Sphinxtrain

We can download the SphinxTrain CMU training package nightly build freely [10]. The execution of SphinxTrain requires additional software too:

- Active Perl: To edit scripts provided by Sphinxtrain [15].

III. TAMIL SPEECH RECOGNITION

Tamil is a Dravidian language spoken predominantly in the state of Tamilnadu in India and Sri Lanka. It is the official language of the Indian state of Tamilnadu and also has official status in Sri Lanka and Singapore. With more than 77 million speakers, Tamil is one of the widely spoken languages of the world. In Tamil, vowels are classified into short (*Kuril*) and long (*Netil*). There are five short vowels and seven long vowels which include two diphthongs. Consonants are classified into three categories with six in each category: hard, soft and medium with a total of 18 consonants. The soft consonants are also known as nasals.

The classification of consonants is based on the place of articulation. The vowels and consonants combine to form 216 compound characters. The compound characters are formed by placing dependent vowel markers on either one side or both sides of the consonant.There is one more special letter *aytham* (∴) used in classical Tamil and rarely found in modern Tamil.Summing up the vowels, consonants and compound letters, there are in total 247 characters in standard Tamil alphabet [17].

IV. STEPS TO CREATE AN APPLICATION

1. Preparation of speech corpus for train the system.
2. Feature Extraction
3. Generation of an Acoustic model.
4. Generation of Language model.
5. System settings.

4.1 Corpus Preparation

An in-house corpus was created from 46 Tamil words. A number of 15 Tamil speakers were asked to utter all digits at one time. Depending on this, the corpus consists of 690 tokens. All the 690 tokens were used for training and testing phases. Recording files must be in MS WAV format with specific sample rate - 16 kHz, 16 bit, mono for desktop application, 8 kHz, 16 bit, mono for telephone applications. Audio files shouldn't be very long and shouldn't be very short. Optimal length is not less than 5 seconds and not more than 30 seconds.

4.2. Feature Extraction

For speech recognition, the most commonly used acoustic features are **Mel-Scale Frequency Cepstral Coefficient** (**MFCC**). MFCC takes human perception sensitivity with respect to frequencies into consideration, and therefore are best for speech recognition. The default for sound files used in Sphinx is a rate of 16 thousand samples per second (16 KHz). If this is the case, the etc/feat.params file will be automatically generated with the recommended values. Recommended values are

-alpha 0.97
-dither yes
-doublebw no
-nfilt 40
-ncep 13
-lowerf 133.33334
-upperf 6855.4976
-nfft 512
-wlen 0.0256

By doing the following steps we can extract features from the speech corpus [4].

1. Pre-emphasis,
2. Frame blocking,
3. Hamming windowing,
4. Fast Fourier Transform or FFT,
5. Triangular Band pass Filters,
6. Discrete cosine transform,
7. Log energy,
8. Delta cepstrum.

4.3 Acoustic Model

An acoustic model represents statistically a range of possible audio representations for the phonemes. For training acoustic models is necessary a set of feature files computed from the audio training data, one each for every recording in the training corpus. Each recording is transformed into a sequence of feature vectors consisting of the Mel-Frequency Cepstral Coefficients (MFCCs).

The training process consists of the following steps: Obtain a corpus of training data and for each utterance, convert the

audio data to a stream of feature vectors, convert the text into a sequence of linear HMMs using the pronunciation lexicon, and find the best state sequence or state alignment through the sentence HMM for the corresponding feature vector sequence. For each senone, gather all the frames in the training corpus that mapped to that senone in the above step and build a suitable statistical model for the corresponding collection of feature vectors. The circularity in this training process is resolved using the iterative Baum- Welch or forward-backward training algorithm [17].

In our application Tamil words the preceding set was used to train the states of the HMM that corresponds to corresponding dictionary (see Table-2) mapping each word to the corresponding symbolic representation, to feed the sphinxTrain application that generated the corresponding Acoustic model for the application. The following table shows an excerpt from the dictionary used for the training and recognition phases:

Table-1: Excerpt from Dictionary of Isolated Tamil Words Application

EQUIVALENT ENGLISH TRANSLITERATE	EQUIVALENT TAMIL WORD
AGATHI	அகத்தி
AVARAIKKAI	அவரைக்காய்
KALAN	காளான்
KUDAMILAGAI	குடைமிளகாய்
MULAIKEERAI	முளைக்கீரை
OMAILLAI	ஓமஇலை
PATTANI	பட்டாணி
PUDALANGA	புடலங்காய்
SUNDAIKAI	சுண்டைக்காய்

Sources: Authors Compilation.

4.4 Language Model

Language model or grammar model which models patterns of word usage. This is normally customized for the application. Every word in the language model must be in the pronunciation dictionary. In isolated speech recognition language model is called as grammar file.

Figure-2: Excerpt from the Grammar File of Tamil Isolated Words Application

```
/**************************************
 * JSGF Grammar for HelloIsolated*
 **************************************/
grammar hello;

public <greet> = ( AGATHI | MULAIKEERAI |
KALAYANAPOOSANIKKAI | KATHIRIKKAI | BEETROOT
| AVARAIKKAI | THULASI | PAWAKKAI | KUDAMILAGAI
| MATTAIKOSU | CAULIFLOWER | CARROT | OMAILLAI |
KOTHAVARANGAI | SEPPANKILANGU |
KOTHAMALLIILAI | THENGAI | SEEMAIKATHIRIKKAI |
VELLIRIKKAI | MURANGAIKKAI | VENDHAYAILAI |
POODINAILAI | FRESHBEANS | PATTANI |
PULICHAIKEERAI | VENDAIKKAI |
YELUMICHAMPALAM | KALAN);
```

4.5 System Settings

An automatic Speech recognizer system like Sphinx4 uses two types of language-dependent models: an acoustic model and a language model or grammar model. In our application, we proceed to the modification of those two elements as described before.

The core recognition process is provided automatically by the sphinx engine using the appropriate language and acoustic models. The sphinx framework must be configured using an xml based configuration file. The algorithms use, the feature extraction algorithms and all other important aspects of any speech recognition application can be customized to tailor the needs of the researcher.

V. RESULT

In order to evaluate the performances of the application, we performed some experiments on different individuals each one of them was asked to utter 46 isolated Tamil words. We recorded the number of words that were correctly recognized, and then a mean recognition ratio for each tester was calculated.

Table-2: Test Results for Individual Speakers

Speaker	Test1	Test2	Test3	Mean Ratio
Speaker1	40	38	42	86.9%
Speaker2	38	40	43	87.6%
Speaker3	42	42	39	89.1%

VI. CONCLUSIONS

This paper talked about Isolated Tamil speech recognition system was designed to investigate the process of automatic speech recognition. This system was based on CMUSphinx-4 from Carnegie Mellon University. An application was presented to demonstrate the possible adaptability of this system to Tamil speech.

VII. REFERENCES

[1] A., Yousfi, "Introduction de la vitesse d'elocution et de l'energie dans un modele de reconnaissance automatique de la parole".

[2] S., Young, (2002). "The HTK hidden Markov model toolkit: Ph.D. thesis, Oujda University, 2002, Design and Philosophy", Cambridge University Engineering Department, UK, Tech. Rep. CUED/FINFENG/TR152, September.

[3] N., Deshmukh; A., Ganapathiraju; J., Hamaker; J., Picone, and M., Ordowski, (1999). "A Public Domain Speech-To-Text System", Proceedings 6th European Conference Speech Communicn and Tehnolgy, Volume: 5, B pest, Hungary, September, pp. 2127-2130.

[4] X., X. Li; Y., Zhao; X., Pi; L., H. Liang, and A., V. Nefian, (2002). "Audio-Visual Continuous Speech Recognition Using A Coupled Hidden Markov Model", Proceeding 7th International Conference Spoken Language Processing, Denver, CO, September, pp. 213-216.

[5] K., F. Lee; H., W. Hon, and R., Reddy, (1990). "An Overview of the Sphinx Speech Recognition System", IEEE Transactions on Acoustics, Speech and Signal Processing, January, Volume: 38, Number: 1, pp. 35-45.

[6] X., Huang; F., Alleva; H., W. Hon; M., Y. Hwang, and R., W2, Rosenfeld, (1993). "The SPHINX-II Speech Recognition System: W3 An Overview", Computer Speech and Language, Volume: 7, Number: 2, pp. 137-148.

[7] M., K. Ravishankar, (1996). "Efficient Algorithms for Speech Recognition", PhD Thesis (CMU Technical Report CS-96-143), Carnegie Mellon University, Pittsburgh, PA.

[8] X., D. Huang; Y., Ariki, and M., A. Jack, (1990). "Hidden Markov Models for Speech Recognition", Edinburgh: Edinburgh University Press, C.

[9] G., Z. Hong, (2002). "Speech Recognition Techniques for Digital Video Library", University of Hong Kong.

[10] Carnegie Mellon University, Sphinx-4, Available: http://cmusphinx.sourceforge.net.

[11] X., D. Huang, (1989). "The SPHINX-II Speech Recognition System: An Overview", Computer Speech and Language, Volume: 2.

[12] K., F. Lee, (1989). "Automatic Speech Recognition the Development of the SPHINX System", Kluwer Academic Publishers.

[13] http://cmusphinx.sourceforge.net/sphinx4.

[14] Sun Microsystems. Available: http://java.sun.com.

[15]http://cmusphinx.sourceforge.net/sphinx4/doc/Eclipse.html , "Sphinx-4 Setup on Eclipse".

[15] http://activestate.com

[16] R., Thangarajan; A., M. Natarajan, and M., Selvam, (2009). "Syllable Modeling in Continuous Speech Recognition for Tamil Language", November.

[17] http://www.speech.kth.se/wavesurfer

*****

DATA INTEGRITY IN CLOUD CATCHING

Deepika Shanmugan[46] Shameena S.[47] Vinothselvin S.[48]

ABSTRACT

Cloud computing has been envisioned as the de-facto solution to the rising storage costs of IT Enterprises. With the high costs of data storage devices as well as the rapid rate at which data is being generated it proves costly for enterprises or individual users to frequently update their hardware. Apart from reduction in storage costs data outsourcing to the cloud also helps in reducing the maintenance. Cloud storage moves the user's data to large data centers, which are remotely located, on which user does not have any control.

However, this unique feature of the cloud poses many new security challenges which need to be clearly understood and resolved. We provide a scheme which gives a proof of data integrity in the cloud which the customer can employ to check the correctness of his data in the cloud. This proof can be agreed upon by both the cloud and the customer and can be incorporated in the Service Level Agreement (SLA).

KEYWORDS

Cloud Computing, Service Level Agreement (SLA), Data, Storage, Security, Integrity, Proof of Retrievability etc.

I. INTRODUCTION

Data outsourcing to cloud storage servers is raising trend among many firms and users owing to its economic advantages. This essentially means that the owner (client) of the data moves its data to a third party cloud storage server which is supposed to - presumably for a fee - faithfully store the data with it and provide it back to the owner whenever required.As data generation is far outpacing data storage it proves costly for small firms to frequently update their hardware whenever additional data is created. It can also assure a reliable storage of important data by keeping multiple copies of the data thereby reducing the chance of losing data by hardware failures. Storing of user data in the cloud despite its advantages has many interesting security concerns which need to be extensively investigated for making it a reliable solution to the problem of avoiding local storage of data.

In this paper we deal with the problem of implementing a protocol for obtaining a proof of data possession in the cloud sometimes referred to as Proof Of Retrievability (POR).This problem tries to obtain and verify a proof that the data that is stored by a user at a remote data storage in the cloud (called cloud storage archives or simply archives) is Not modified by the archive and thereby the integrity of the data is assured. Such verification systems prevent the cloud storage archives from misrepresenting or modifying the data stored at it without the consent of the data owner by using frequent checks on the storage archives. Such checks must allow the data owner to efficiently, frequently, quickly and securely verify that the cloud archive is not cheating the owner. Cheating in this context means that the storage archive might delete some of the data or may modify some of the data.

II. PROJECT DESCRIPTION

2.1 Problem Definition

Storing of user data in the cloud despite its advantages has many interesting security concerns which need to be extensively investigated for making it a reliable solution to the problem of avoiding local storage of data. Many problems like data authentication and integrity (i.e., how to efficiently and securely ensure that the cloud storage server returns correct and complete results in response to its clients' queries, outsourcing encrypted data and associated difficult problems dealing with querying over encrypted domain were discussed in research literature.

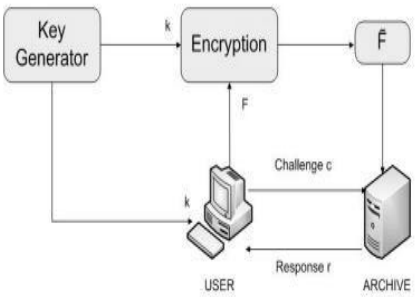
2.2 Overview of Project

Data outsourcing to cloud storage servers is raising trend among many firms and users owing to its economic advantages. This essentially means that the owner (client) of the data moves its data to a third party cloud storage server which is supposed to - presumably for a fee - faithfully store the data with it and provide it back to the owner whenever required. Purpose of developing proofs for data possession at untrusted cloud storage servers we are often limited by the resources at the cloud server as well as at the client. Given that the data sizes are large and are stored at remote servers, accessing the entire file can be expensive in I/O costs to the storage server.

Also transmitting the file across the network to the client can consume heavy bandwidths. Since growth in storage capacity has far outpaced the growth in data access as well as network bandwidth, accessing and transmitting the entire archive even occasionally greatly limits the scalability of the network resources. Furthermore, the I/O to establish the data proof interferes with the on-demand bandwidth of the server used for normal storage and retrieving purpose.

2.2.1 System Architecture

Figure-1: Schematic View of a Proof of Retrievability Based on Inserting Random Sentinels In Data File



Sources: Authors Compilation.

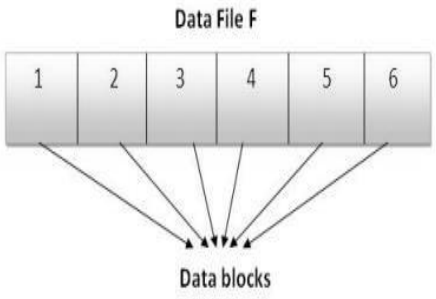
[46] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, deepikash90@gmail.com
[47] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, shamib.tech12@gmail.com
[48] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, vinothselvin@gmail.com

2.2.2 Metadata Generation

A. Setup Phase

Let the verifier V wishes to the store the file F with the archive. Let this file F consist of n file blocks. We initially preprocess the file and create metadata to be appended to thefile. Let each of the n data blocks have m bits in them. Atypical data file F which the client wishes to store in the cloud is shown in Figure-2.The initial setup phase can be described in the following steps:

Figure-2: Data File F with 6 Data Blocks



Sources: Authors Compilation.

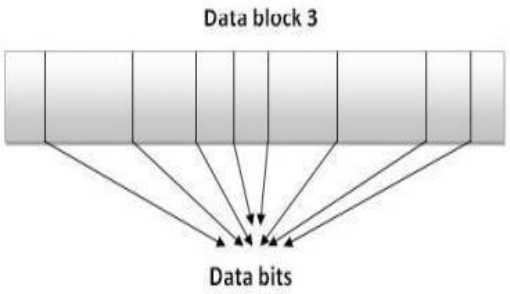
1) Generation of Meta-Data

Let g be a function defined as follows: g(i, j) → {1..m}, i ∈ {1..n}, j ∈ {1..k} (1)

Where k is the number of bits per data block which we wish to read as meta data. The function g generates for each data block a set of k bit positions within the m bits that are in the data block. Hence, g(i, j) gives this data block.

The value of k is in the choice of the verifier and is secret known only to him. Therefore for each data block we get a set of k bits and in total for all the n blocks we gent ∗ k bits. Let mi represent the k bits of meta data for the it block. Figure-3 shows a data block of the file F with random bits selected using the function g.

Figure-3: Data Block of File F with Random Bits Selected In It



Sources: Authors Compilation.

2) Encrypting the Meta Data

Each of the meta data from the data blocks mi is encrypted by using a suitable algorithm to give a new modified meta data MiWithout loss of generality we show this process by using simple XOR operation. Let h be a function which generates k bit integer αi for each i. This function is a secret and is known only to the verifier V.h : i → αi, αi ∈ {0..2n} (2)
For meta data (mi) of each data block the number is added to get a new k bit number Mi,

Mi = mi + αi (3)

In this way we get a set of n new meta data bit blocks. The encryption method can be improvised to provide still stronger protection for verifiers data.

3) Appending of Meta Data

All the meta data bit blocks that are generated using the above procedure are to be concatenated together. This concatenated meta data should beappended to the file F before storing it at the cloud server. Thefile F along with the appended meta data Fe is archived with the cloud. Figure-4 shows the encrypted file Fe after appending the meta data to the data file F.

Figure-4: Encrypted file 'F' Stored in Cloud

Sources: Authors Compilation.

B. Verification Phase

Let the verifier V want to verify the integrity of the file F.It throws a challenge to the archive and asks it to respond. The challenge and the response are compared and the verifieraccepts or rejects the integrity proof. Suppose the verifier wishes to check the integrity of matchlock. The verifier challenges the cloud storage server by specifying the block number i and a bit number j generated by using the function g which only the verifier knows.

Thieveries also species the position at which the meta data Corresponding the block i is appended. This meta data will be a k-bit number. Hence the cloud storage server is required to send k+1 bits for verification by the client.The meta data sent by the cloud is decrypted by using the number αi and the corresponding bit in this decrypted metadata is compared with the bit that is sent by the cloud. Any mismatch between the two would mean a loss of the integrity of the client's data at the cloud storage.

2.3 Module Discription

2.3.1 Cloud Storage

Data outsourcing to cloud storage servers is raising trend among many firms and users owing to its economic advantages. This essentially means that the owner (client) of the data moves its data to a third party cloud storage server

which is supposed to - presumably for a fee - faithfully store the data with it and provide it back to the owner whenever required.

2.3.2 Simple Archives

This problem tries to obtain and verify a proof that the data that is stored by a user at remote data storage in the cloud (called cloud storage archives or simply archives) is not modified by the archive and thereby the integrity of the data is assured.

Cloud archive is not cheating the owner, if cheating, in this context, means that the storage archive might delete some of the data or may modify some of the data. While developing proofs for data possession at untrusted cloud storage servers we are often limited by the resources at the cloud server as well as at the client.

2.3.3 Sentinels

In this scheme, unlike in the key-hash approach scheme, only a single key can be used irrespective of the size of the file or the number of files whose retrievability it wants to verify. Also the archive needs to access only a small portion of the file F unlike in the key-has scheme which required the archive to process the entire file F for each protocol verification. If the prover has modified or deleted a substantial portion of F, then with high probability it will also have suppressed a number of sentinels.

2.3.4 Verification Phase

The verifier before storing the file at the archive preprocesses the file and appends some Meta data to the file and stores at the archive. At the time of verification the verifier uses this Meta data to verify the integrity of the data. It is important to note that our proof of data integrity protocol just checks the integrity of data i.e. if the data has been illegally modified or deleted. It does not prevent the archive from modifying the data.

III. CONCLUSION AND FUTURE WORKS

In this paper we have worked to facilitate the client ingetting a proof of integrity of the data which he wishes to store in the cloud storage servers with bare minimum costs and efforts. Our scheme was developed to reduce the computational and storage overhead of the client as well as to minimize the computational overhead of the cloud storage server.

We also minimized the size of the proof of data integrity so as to reduce the network bandwidth consumption. At the client we only store two functions, the bit generator function g, and the function h which is used for encrypting the data. Hence the storage at the client is very much minimal compared to all other schemes [4] that were developed. Hence this scheme proves advantageous to thin clients like PDAs and mobile phones.

The operation of encryption of data generally consumes a large computational power. In our scheme the encrypting process is very much limited to only a fraction of the whole data thereby saving on the computational time of the client. Many of the schemes proposed earlier require the archive to perform tasks that need a lot of computational power to generate the proof of data integrity [3]. But in our scheme the archive just need to fetch and send few bits of data to the client.

The network bandwidth is also minimized as the size of the proof is comparatively very less(k+1 bits for one proof).It should be noted that our scheme applies only to static storage of data. It cannot handle to case when the data need to be dynamically changed.

Hence developing on this will be a future challenge. Also the number of queries that can be asked by the client is fixed apriori. But this number is quite large and can be sufficient if the period of data storage is short. It will be a challenge to increase the number of queries using this scheme.

IV. REFERENCES

[1] E., Mykletun; M., Narasimha, and G., Tsudik, (2006). "Authentication and Integrity in Outsourced Databases", Trans. Storage, Volume: 2, Number: 2, pp. 107–138.

[2] D., X. Song; D., Wagner, and A., Perrig, (2000). "Practical Techniques for Searches on Encrypted Data", SP '00: Proceedings of 2000 IEEE Symposium on Security and Privacy. Washington, DC, USA: IEEE Computer Society, 2000, pp. 44.

[3] A., Juels, and B., S. Kaliski, Jr., (2007). "Pors: Proofs of Retrievability for Large Files", CCS '07: Proceedings of the 14th ACM conference on Computer and Communications Security, New York, NY, USA: ACM, pp. 584–597.

[4] G., Ateniese; R., Burns; R., Curtmola; J., Herring; L., Kissner; Z., Peterson, and D., Song, (2007). "Provable Data Possession at Untrusted Stores", CCS '07: Proceedings of the 14th ACM conference on Computer and Communications Security, New York, NY, USA: ACM, pp. 598–609.

*****

MODELING OF DILUTION IN PLASMA TRANSFERRED ARC HARDFACING

K. Mohan[49] N. Murugan[50]
T. Kannan[51] G. Krishna Kumar[52]

ABSTRACT

Hardfacing is a form of surfacing that is applied for the purpose of increasing wear, abrasion, and erosion resistance properties. This process is widely used to repair the railway rolling stock, earth moving and agricultural machineries. In hardfacing, the most important aspect is dilution of filler metal with the substrate due to arc penetration. The composition and properties of hardfaced components are strongly influenced by dilution. Hence, control of dilution is essential to achieve the required wear resistance properties. The amount of dilution will be vitally affected by welding process, technique, chemical composition of weld metal and substrate used. This paper presents about an experimental study carried out to model the effect of various parameters of plasma transferred arc welding in tungsten carbide hardfacing of 316L stainless steel plates on dilution.

KEYWORDS

Hardfacing, Dilution, Mathematical Model, Tungsten Carbide, Stainless Steel etc.

I. INTRODUCTION

Hardfacing is a process of depositing a filler metal on a substrate to impart wear and erosion resistance properties to the surface that is not intrinsic to the underlying substrate. Hardfacing is applied for the purpose of increasing wear, abrasion, and erosion resistance properties. This process is widely used to repair the railway rolling stock, earth moving and agricultural machineries [1]. Hardfacing is carried out by oxyfuel welding, gas tungsten arc welding, gas metal arc welding, shielded metal arc welding and flux cored arc welding. Plasma transferred arc hardfacing and laser beam hardfacing are relative newcomers to this arena. Due to its various advantages such as low dilution, high deposition rate, etc Plasma transferred arc hardfacing has became a popular process especially for depositing stellite and tungsten carbide powders.

In hardfacing process, the most important aspect is dilution. The composition and properties of hardfacing metals are strongly influenced by dilution. Hence, control of dilution is essential to achieve the required wear resistance properties. The main target in each hardfacing process is to maintain the dilution at minimum level. Dilution can be effectively controlled by proper selection of process parameters after thoroughly studying the direct and interaction effects of process parameters on dilution [2]. The amount of dilution will be vitally affected by the welding process, technique, chemical composition of weld metal and substrate used [1]. A value between 10 and 15% is generally considered optimum [1]. Although each process has an expected dilution factor, it is essential to establish a standard procedure to obtain low dilution, good bonding and a sound deposit. Welding procedures are often established by trial and error method. Often a satisfactory working point can be chosen on the basis of combined experimentation, experience and existing data. But it is essential to show that such a working point indeed tolerant and reproducible, and a quantitative estimate of variability in the manufactured quality has to be provided for practical purposes. This can be easily achieved with the help of development of a mathematical model.

This paper presents about an experimental study carried out to model the effect of various parameters of plasma transferred arc welding in tungsten carbide hardfacing of 316L stainless steel plates. The experiments were conducted based on five-factor five level central composite rotatable designs. Mathematical model was developed to predict dilution. The developed mathematical model has been checked for their adequacy and significance.

II. EXPERIMENTAL DETAILS

The experiments were conducted using an automatic PTA hardfacing system. The experimental setup of Plasma Transferred Arc Welding process used for this work is shown in Figure-1.

Substrate Used: Stainless steel plate designated as 316L was used for this study. The chemical composition of this Stainless steel is given in Table-1.

Metal Powder Used: Tungsten carbide designated as Durmat-61PTA, was used for depositing the beads. The chemical composition of the powder is given in Table-1.

Shielding Gas Used: Argon gas was used for shielding.

Table-1: Chemical Composition of Base Metal and Metal Powder

Material	Elements, wt. %										
	C	Si	Mn	P	S	B	Cr	Mo	Fe	WSC	Ni
316L	0.03	0.75	2.00	0.045	0.03	-	18.00	2.00	-	-	10.00
Tungsten carbide	0.02	3.00	-	-	-	3.00	-	-	< 2.0	60	Balance

Sources: Authors Compilation.

Table-2: Limits & Levels of Welding Parameters

Parameter	Unit	Notation	Factor Levels				
			- 2	- 1	0	+ 1	+ 2
Welding current	Amps	I	135	142	149	156	163
Welding speed	mm/min	S	77	81	85	89	93
Powder feed rate	g/min	F	34	36	38	40	42
Oscillation frequency	Cycles/min	H	36	38	40	42	44
Stand off	mm	N	8	9	10	11	12

Sources: Authors Compilation.

[49]Professor, K. S. Rangasamy College of Technology, Tamil Nadu, India, mohanmechksr@yahoo.com
[50]Professor, Coimbatore Institute of Technology Tamil Nadu, India, drmurugan@hotmail.com
[51]Principal, SVS College of Engineering, Tamil Nadu, India, kannan_svsce@yahoo.com
[52] Assistant Professor, SVS College of Engineering, Tamil Nadu, India, krishnakumark555@gmail.com

Figure-1: Experimental Setup

Sources: Authors Compilation.

III. EXPERIMENTAL PROCEDURES

The step-by-step experimental procedures involved in this study are briefly explained below.

3.1 Identification of Process Parameters

The independently controllable process parameters identified to carry out experimental work were welding current (A), travel speed (S), powder feed rate (F), oscillation frequency (H) and standoff (N).

3.2 Identification of Response Variables

The chosen response was percentage dilution (D). The identified process parameters are Welding current, Welding speed, Oscillation Frequency, Powder Feed Rate, Stand-Off.

3.3 Finding Limits and Levels of the Process Parameters

The working ranges of all selected parameters were fixed by conducting trial runs. This was carried out by varying one of the parameters while keeping the rest of them at constant values [3]. The working range of each process parameters was decided upon by inspecting the weld bead for a smooth appearance without any visible defects such as surface porosity, undercut, etc. [4]. The upper limit of a factor is coded as +2 and the lower limit is coded as -2. The coded values for intermediate values are calculated using the following equation 1 [4].

$$X_i = 2\,[2X - (X_{max} + X_{min})] / (X_{max} - X_{min})] \text{----------(1)}$$

Where X_i is the required coded value of a variable X, X is any value of the variable from X_{min} to X_{max}, X_{min} is the lower limit of the variable and X_{max} is the upper limit of the variable. The values of other secondary parameters that were kept constant during the trial runs are given below:

Base plate (Substrate) thickness: 25 mm
Shielding gas flow rate : 12L / min
Tungsten carbide particle size : -160 to + 63µm
Electrode polarity : Positive
Preheating temperature : 500° C

The decided limits and levels of the selected process parameters with their units and notations are given in Table-2.

3.4 Development of Design Matrix

The design matrix chosen to conduct the experiment was a central composite rotatable design [5].

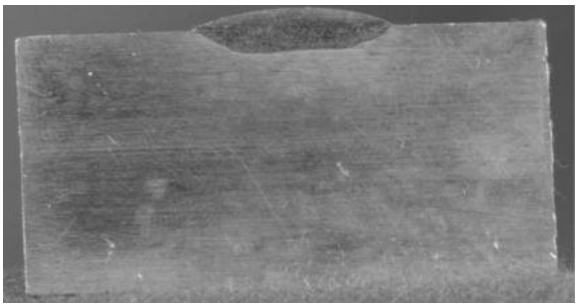
3.5 Conducting Experiments as per the Design Matrix

Test pieces of size 100 mm X 76 mm X 25 mm were prepared and preheated. The experiments were conducted by using wave bead technique. Adopting the same procedure, 32 experimental runs were conducted as per the design matrix at random.

3.6 Recording of Responses

To measure the percentage of dilution, the hardfaced plates were cross-sectioned. The sectioned specimen end faces were polished and etched using 2% nital solution to display weld bead dimensions [6]. Bead profiles were traced using a optical profile projector at a magnification of X10. Bead configuration of a typical hardfaced specimen is shown in Figure-2. The areas of base metal melted and weld metal forming reinforcement were measured with help of a digital planimeter and percentage of dilution was calculated. The calculated dilutions are presented in Table-3. Using these values, mathematical regression model was developed as follows:

Figure-2: Bead Configuration of a Typical Hard Faced Specimen



Sources: Authors Compilation.

Table-3: Design Matrix and Observed Values of Dilution

Trial No.	Design Matrix					Dilution
	I	S	F	H	N	D (%)
01	-1	-1	-1	-1	1	26.95
02	1	-1	-1	-1	-1	31.60
03	-1	1	-1	-1	-1	30.45
04	1	1	-1	-1	1	29.60
05	-1	-1	1	-1	-1	31.35
06	1	-1	1	-1	1	27.80
07	-1	1	1	-1	1	29.90
08	1	1	1	-1	-1	32.55
09	-1	-1	-1	1	-1	28.40
10	1	-1	-1	1	1	35.75
11	-1	1	-1	1	1	24.60
12	1	1	-1	1	-1	35.80
13	-1	-1	1	1	1	32.70
14	1	-1	1	1	-1	24.70
15	-1	1	1	1	-1	36.55
16	1	1	1	1	1	34.55
17	-2	0	0	0	0	30.00
18	2	0	0	0	0	25.60
19	0	-2	0	0	0	28.85
20	0	2	0	0	0	33.50
21	0	0	-2	0	0	31.40
22	0	0	2	0	0	29.10
23	0	0	0	-2	0	28.35
24	0	0	0	2	0	24.40
25	0	0	0	0	-2	29.85
26	0	0	0	0	2	29.75
27	0	0	0	0	0	22.25
28	0	0	0	0	0	24.75
29	0	0	0	0	0	23.80
30	0	0	0	0	0	27.80
31	0	0	0	0	0	32.55
32	0	0	0	0	0	27.90

Sources: Analysis.

3.7 Development of Mathematical Model

The response function representing dilution can be expressed [5] using the equation 2.

$Y = f(X_1, X_2, X_3, X_4, X_5)$ ---------- (2)
Where, Y = response [i.e., dilution]

X_1 = Welding current (I), Amps
X_2 = Welding speed (S), mm/min
X_3 = Powder feed rate (F), g/min
X_4 = Oscillation frequency (H), Hertz and
X_5 = Stand off (N), mm

The second order response surface model [6] for five selected factors is given by the equation 3:

$$Y = \beta_o + \sum_{i=1}^{5} \beta_i X_i + \sum_{i=1}^{5} \beta_{ii} X_i^2 + \sum_{\substack{i=1 \\ i<j}}^{5} \beta_{ij} X_i X_j \text{ ------- (3)}$$

The above second order response surface model could be expressed as follows (equation 4)

$Y = \beta_o + \beta_1 A + \beta_2 S + \beta_3 F + \beta_4 H + \beta_5 N + \beta_{11} A^2 + \beta_{22} S^2 + \beta_{33} F^2 + \beta_{44} H^2 + \beta_{55} N^2 + \beta_{12} AS + \beta_{13} AF + \beta_{14} AH + \beta_{15} AN + \beta_{23} SF + \beta_{24} SH + \beta_{34} FH + \beta_{35} FN + \beta_{45} HN$ ---------- (4)

Where β_0 is the free term of the regression equation, the coefficients β_1, β_2, β_3 β_4 and β_5 are linear terms, the coefficients β_{11}, β_{22}, β_{33}, β_{44} and β_{55} are the quadratic terms, and the coefficients β_{12}, β_{13}, β_{14} , β_{15}, β_{23} , β_{24} , β_{34}, β_{35} and β_{45} are the interaction terms. The values of the coefficients of the polynomial equation 3.4 were calculated [7] using QA six-sigma software mathematical model constructed using the calculated coefficients are given below:

Percentage dilution (D) = 26.307 + 0.110I +1.002S + 0.098F + 0.206H - 0.406N + $0.524I^2$ +$1.368S^2$ + $1.136F^2$ + $0.168H^2$ +$1.024N^2$ + 0.659IS - 2.078IF + 0.353IH + 0.978IN +1.203SF + 0.322SH -1.491SN + 0.059FH + 0.572FN + 0.866HN

3.8 Checking Adequacy of the Developed Model

The adequacy of the developed model was tested using analysis of variance (ANOVA) technique [7]. The analysis of variance for the response is presented in Table 4. From table, it is found that the model is adequate.

Table-4: Analysis of Variance for Testing Adequacies of the Models

Parameter	1st order terms		2nd order terms		Lack of fit		Error terms		F-ratio	R-ratio	Remarks
	SS	DF	SS	DF	SS	DF	SS	DF			
% Dilution	29.605	5	279.799	15	47.499	6	68.667	5	0.594	4.677	adequate

Sources: Calculated Data.

F-ratio (6, 5, 0.05) = 4.95, SS-Sum of squares
R-ratio (20, 5, 0.05) = 4.56, DF-Degrees of freedom

3.9 Development of Final Model

After elimination of insignificant coefficients, final regression model was constructed using significant coefficients and is as follows:

Percentage dilution (D) = 26.946 + 0.110I +1.002S + 0.098F+ 0.206H - 0.406N + $1.314S^2$ +$1.083F^2$ + $0.971N^2$ – 2.078IF + 0.978IN + 1.203SF -1.491SN + 0.866HN.

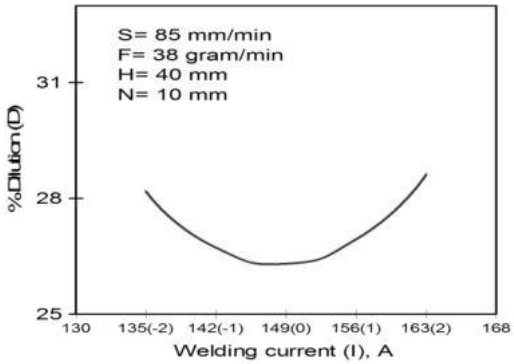
IV. RESULTS AND DISCUSSIONS

The developed model can be used to predict the dilution by substituting the coded values of the respective process parameters. The responses calculated from the model for each set of coded welding parameters are represented in graphical form in Figures 3 to 7.

4.1 Effect of Welding Current on Dilution

Figure-3 shows that dilution initially decrease to a minimum with increase in welding current and then increases steadily with the further increase in welding current.

Figure-3

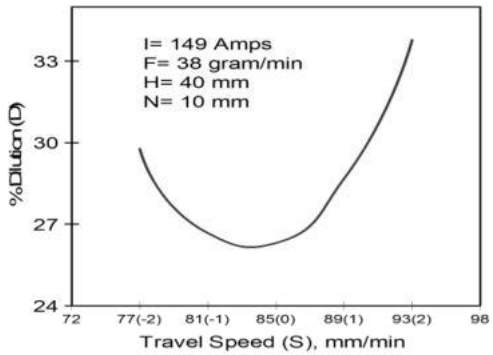


Sources: Authors Compilation.

4.2 Effect of Welding Speed on Dilution

From Figure-4, it is evident that dilution initially decreases with increase in welding speed. At lower welding speed the weld pool cushions the effect of arc and prevents deeper penetration [10] resulting in lower dilution.

Figure-4



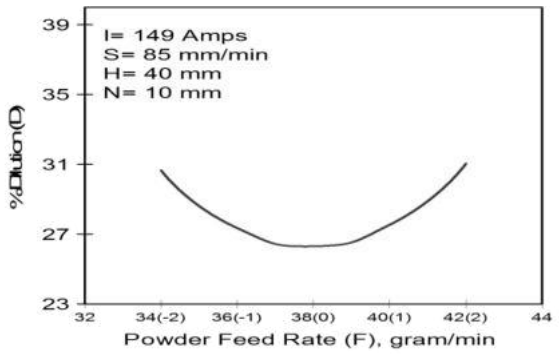
Sources: Authors Compilation.

4.3 Effect of Powder Feed Rate on Dilution

From Figure-5, it is evident that dilution decrease initially with increase in powder feed rate but increases further with the increase in powder feed rate. From figure, following inferences can be obtained: (i) at lower and higher powder feed rate, the dilution is greater when compared to intermediate level of powder feed rate. At lower powder feed rate, more percentage of heat generation is utilized for melting powders, and a lower percentage of heat generation is used for melting the substrate. Due to this condition, amount of base metal melted is lower, and hence the dilution is lower at the lower powder feed rate. On other hand, the substrate melted is lower due to less amount

of heat is utilized for this purpose; the excess amount of molten metal produces deeper penetration [7, 11]. This may be the reason for an increase in dilution at higher powder feed rate.

Figure-5

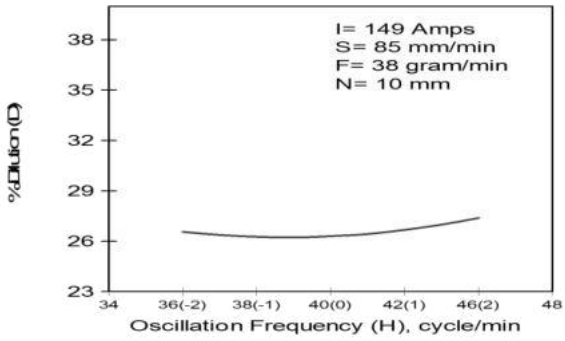


Sources: Authors Compilation.

4.4 Effect of Torch Oscillation Frequency on Dilution

It is clear from Figure-6 that dilution decrease initially with increase in torch oscillation frequency to a minimum value and then increases with the increase in torch oscillation frequency. From the figure, the following inferences can be obtained: (i) if torch oscillation frequency is higher, then the dilution is slightly higher and vice-versa. At higher torch oscillation frequency, the torch weaving speed is greater and covers wider region of substrate material results in higher dilution.

Figure-6

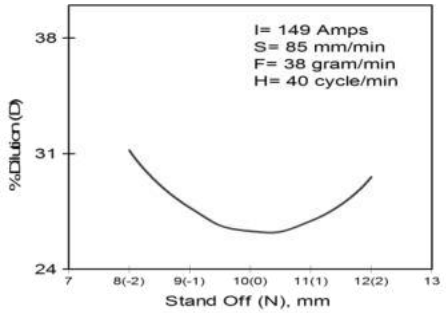


Sources: Authors Compilation.

4.5 Effect of Stand-off on Dilution

It is evident from Figure-7 that dilution decrease slightly with increase in stand-off. Increase in stand-off increases the circuit resistance, which reduces the welding current. This decrease of welding current reduces penetration of arc and hence reduces dilution [12]. Further increase in stand-off distance, voltage may increase resulting in more heat input & dilution increases.

Figure-7



Sources: Authors Compilation.

V. CONCLUSIONS

Percentage dilution initially slightly decreases to a minimum with increase in welding current and then increases steadily with the further increase in welding current. Percentage dilution initially slightly decreases to a minimum with increase in travel speed and then increases steadily with the further increase in travel speed.

At lower and higher powder feed rate, the percentage dilution is greater when compared to intermediate level of powder feed rate. Torch frequency has less significant influence on percentage of dilution is higher and vice-versa. Percentage dilution decreases with increase in stand off.

VI. REFERENCES

1. Horsfield, A. M., (1980). "Weld Surfacing and Hardfacing", The Welding Institute, Abington Hall, pp. 32 - 39

2. ASM Handbook (1993), Welding, Brazing and Hardfacing, Vol. 6, USA.

3. Gupta, V. K., and Parmar, R. S., (1989). "Fractional Factorial Technique to Predict Dimensions of the Weld Bead in Automatic Submerged Arc Welding", Journal of Institution of Engineering (India), 70, pp. 67

4. Gunaraj, V., and Murugan, N., (1999). "Application of Response Surface Methodology for Predicting Weld Bead Quality in Submerged Arc Welding of Pipes", Journal of Material Processing Technology, 88:266 – 275.

5. Cochran, W. G., and Cox, G. M., (1957) "Experimental Designs", Second Edn., New York: John Wiley & Sons.

6. Harris, P., and Smith, B. L., (1983). "Factorial Technique for Weld Quality Prediction", Metal Construction, 15(11):661-666.

7. Douglas C., Montgomery. (2003), "Design and Analysis of Experiments", 5th edn., John Wiley & Sons (ASIA) Pte Ltd.

8. Jean, Cornu, (1988), "Advanced Welding Systems", Volume: 2, IFS (Publications) Ltd., UK.

9. Balasubramanian, V.; Lakshminarayannan, A. K.; Varahamoorthy, R., and Babu, S., (2008). "Understanding the Prameters Controlling Plasma Transferred arc Hardfacing Using Response Surface Methodology", Materials & Manufacturing Processes, 23: pp. 1-9.

10. Kim, J. W., and Na, S. J., (1995). "A study on the Effect of Contact Tube-To-Work Piece Distance On Weld Pool Shape In Gas Metal Arc Welding", Welding Journal, 74 (5):141s-152s.

*****

CLOUD BASED VIRTUAL LAB

Joshi S.[53] M. C. Saravanan[54]

ABSTRACT

Most of the developers with innovative ideas for new internet era can't express their ideas and extend their ideas by large capital outlays in software and maintenance. Many Institution also experience this hurdle that to invest the large amount. Some concern can't able to buy the supporting Software and some feel lazy to install and configure the Software. These outcomes are analyzed and as a result, the project "Cloud based Virtual Lab" is to be developed.

This paper deals with the solution based Software as a service platform. Any user, who wants to utilize the software, can call upon the Cloud using Browser architecture and be able to access the software alone and there is no need of vast amount of Money and High Local Storage. We propose the idea of Virtual Lab Environment that provides the Platform as a service and also provides software as a service to the user. It can avoid the high investment and software installation time; the user can access the desired software through URL and can edit, store and save their workouts.

KEYWORDS

Cloud, Virtual Lab, Software, Service, Platform, Browser etc.

I. INTRODUCTION

Cloud computing is a nascent business and technology concept with different meanings for different people.

For application and IT users, it's IT as a service (ITaaS) is, delivery of computing, storage, and applications over the Internet from centralized data centers [3, 2].

For Internet application developers, it's an Internet-scale software development platform and runtime environment.

For infrastructure providers and administrators, it's the massive, distributed data center infrastructure connected by IP networks. In Simple Word, we define the Cloud is the web based computer where we can access the application, store the data and access the data via internet. We use virtualization Technology to realize the Inter cloud architecture because existing cluster resource managers relying on Virtual Environment can give us building blocks such as availability information required to create virtual Execution environments.

In addition, relying on virtualization technologies has also facilitated the realization of cloud computing services. Cloud computing includes three kinds of Internet-accessible services: software-as-a-service (SaaS), platform-as-a-service (PaaS), and infrastructure-as-a-service (IaaS). Here, we consider only IaaS, which aims to provide computing resources or storage as a service to users.

II. BACKGROUND

A) Virtualization

Virtualization is the most profound change that PCs and servers have experienced, said Simon Crosby, chief technology officer for Citrix Systems' Data Center and Cloud Division [1]. "IT departments have long been at the mercy of the technical demands of legacy applications", explained Chris Van Dyke, [3] Microsoft's chief technology strategist for the oil and gas industry. "Now, rather than having to maintain older operating systems because of the needs of a legacy application, IT departments can take advantage of the performance and security gains in a new OS (in one virtual machine) while supporting legacy applications in another. Also, the process of deploying applications becomes simpler, because applications can be virtualized and deployed as a single virtual machine".

Virtualization technology lets a single PC or server simultaneously run multiple operating systems or multiple sessions of a single OS. This lets users put numerous applications—even those that run on different operating systems—on a single PC or server instead of having to host them on separate machines as in the past. The approach is thus becoming a common way for businesses and individuals to optimize their hardware usage by maximizing the number and kinds of jobs a single CPU can handle.

B) Infrastructure as a Service

Today, Many of Cloud Providers offer Services at Various layers of the software stack. At Lower layers, Infrastructure as a Service (IaaS) providers Such as Amazon, Flexi scale and Go Grid allow their Customers to have access to entire Virtual Machines (VMS) hosted by the Provider. IaaS Provide the Whole infrastructure and using this we can able to use the application within the low configuration.

C) Software as a Service

The Higher Layers, Software as a Service (Saas) such as Google Apps, Google docs and Zoho offer complete online applications that can be directly executed by the users. With the help of single email id, we can able to access the service, edit, run and save in the Google docs and Zoho.

D) Platform as a Service

The Providers like g.ho.st, eyeos, icloud offer the platform in which the user can access the virtual operating system and that platform consist of all the services that the local os have. The Providers won't offer the compiler and the executable software.

III. RELATED WORK

Most Web OSs, such as eyeOS and Desktoptwo, feature APIs that let programmers write new programs for the operating system. Sapotek sponsors the Sapodesk open source software community. About 50 projects are under way to develop desktop-like customer-relationship-management, calculator,

[53] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, jehovahjoshi@gmail.com
[54] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, saravananchellan1@gmail.com

and other Web-based applications, including those that work with the company's two Web OSs, noted Sapotek CEO Joshua Rand [2]. Some Web OSs, such as the SGD,deploy standard Windows or Unix applications online via servers that run the programs in virtual machines that handle them as if they were in their native OS.

1) G.ho.st works with various Web applications already on the Internet, such as Zoho for word-processing, spreadsheet, presentation, and other office-productivity functions.

2) Zimbra for calendaring and messaging; Pandora for playing music; and Flickr for displaying and managing photos, noted G.ho.st CEO Zvi Schreiber[1].

3) The Burton Group's Creese said Microsoft's release of Silver light - a cross-browser, cross-platform plug-in that allows development of rich Web applications - could encourage the many programmers familiar with Microsoft-based development tools to also begin creating software for Web OSs. Web OSs generally comes with a limited number of applications. And, the Burton Group's Creese said, the applications might not have features that users like in their traditional programs. Also, users may be too unfamiliar with the software to want to use it.

IV PROPOSED WORK

A. Considerations

As stated earlier, Web operating systems have the features and functionality of traditional desktop operating system. However, Web operating systems typically transfer applications to web server where user can manage his resources through virtual desktop using web browser. At the start of our research we had three main interests which we tried to satisfy.

- Moving from fully personalized familiar desktop on PC to a virtual Lab for the academic institution, is a hard task, as Students will accept nothing less than traditional desktop which they have been adopt for the local desktop. Thus, user data, preferences as well as sessions must be maintained Ensuring that user will always has a personal experience that resembles his fully personalized traditional pc environment.

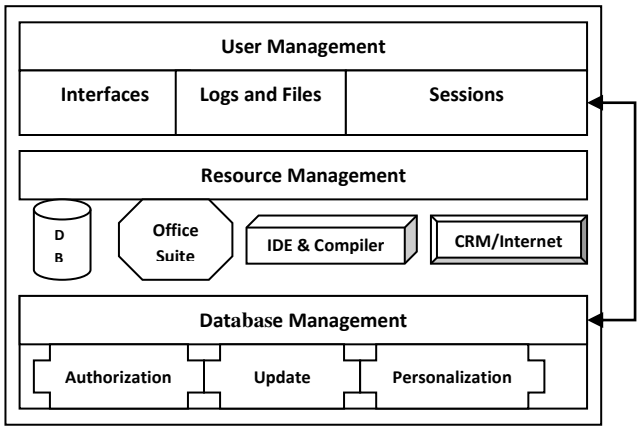
- Semantic web technology plays a significant role in today's web as well as desktop systems. That's why we thought that it was only a matter of time before semantic web techniques thrust in the research of web based lab and operating system.

- A service-oriented architecture (SOA) is seen as the next Evolutionary step in building web-based applications as it provides a set of principles of governing concepts used during phases of systems development. As in n-tier architectures SOA separates presentation/applications, services and data into layers preventing dependency between layers [8, 9, and 10].

Using the considerations of these we proposed a virtual lab. This virtual lab includes compiler, office suites, Database tools, Case tools lab, CRM and the internet applications.

B) Architecture of our Virtual Lab

Figure-1: Architecture



Sources: Authors Compilation.

User Management

In User Management Layer, interaction between the human and the system application is maintained. Some of its functions include the creation of new users in the system, listing of all users, delete users, making interaction between the application and the user etc. There are three main components in this layer. They are Interfaces, Logs and files and Sessions. An interface manages the implementation of the remaining components that form the "engine" like structure. In our project AJAX plays the role of engine and using the AJAX Engine the web applications can retrieve data from the server asynchronously in the background without interfering with the display and behavior of the existing page.

Resource Management

There are various resources which include office suites, database, programming ide, compilers, ERP, CRM. These resources are managed by this resource management layer. All the resources in the lab are divided in five big categories like Office, Network, Database, Programming or Places. Using these resources a student can able to develop the programs, share it, edit and run there without the local installation.

Database Management

In this layer, all the data are managed. The data includes user details, resource details, session details, interaction details. Authorization is provided as we are managing the user details. The database is updated in each session and personalization also managed.

C) Technical Layout

This cloud based virtual lab is mainly design using the embedding application over the WebTop. A web desktop or WebTop is a desktop environment embedded in a web browser or similar client application [12]. A WebTop integrates web applications, web services, client–server applications, application servers, and applications on the local client into a desktop environment using the desktop metaphor. Web desktops provide an environment similar to that of Windows, Mac, or a graphical user interface on UNIX and Linux systems.

It is a virtual desktop running in a web browser. In a WebTop the applications, data, files, configuration, settings, and access privileges reside remotely over the network. Much of the computing takes place remotely. The browser is primarily used for display and input purposes. This was achieved by using the PHP, XML and Java Script (in short AJAX is used to mounting and made communication between client and server).

Ajax Engine acts as a microkernel and the interaction between the client and the cloud. Each core part of the desktop is its own application, using JavaScript to send server commands as the user interacts. As actions are performed using Ajax (such as launching an application), it sends event information to the server. The server then sends back tasks for the client to do in XML format, such as drawing a widget [9, 13].

On the server, Virtual lab uses XML files to store information. This makes it simple for a user to set up on the server, as it requires zero configurations other than the account information for the first user, making it simple to deploy. To avoid bottlenecks that flat files present, each user's information and settings are stored in different files, preventing resource starvation.

We provide various types of applications in this Virtual lab such as IBBoard, IBMessages, IBNav, IBEdit, IBWebsql, IBDeveloper, IBFeed, Moodle, and IBCalendar:

i) IBBoard: It is a tag board program. Each user can add new text, messages, events and superuser can empty the whole log. A header shows the time and name of the user who posted each message. Using this students and staffs can share the timings, sessions and classes.

ii) IBMessages: It allows you to send messages to other system users, as well as read messages other users have sent to you. Although, it does not allow you to send standard e-mails to people outside of the virtual lab. It help to share the links to the off list users too.

iii) IBNav: It is the web browser that launch inside the lab and we can able to access the world wide information and the locale repository too.

iv) IBEdit: Normal editor functions are performed using this application and we can export this edit application to other format like doc, docx, odt to the local system.

v) IBWebsql: This created as a simple replacement for desktop mysql client applications, and to assist staffs and students of mysql for doing queries on their database using a simple interface.

vi) IBDeveloper: It is an Online Integrated Development Environment (IDE). Users would be designing computer Programs such as C, C++, Java, Python using this environment and can able to connect the exported coding with the online websql that we provide.

vii) IBFeed: IBFeed is a simple RSS Feed which gives the news periodically. Using this students can able to update the current news in and around the college.

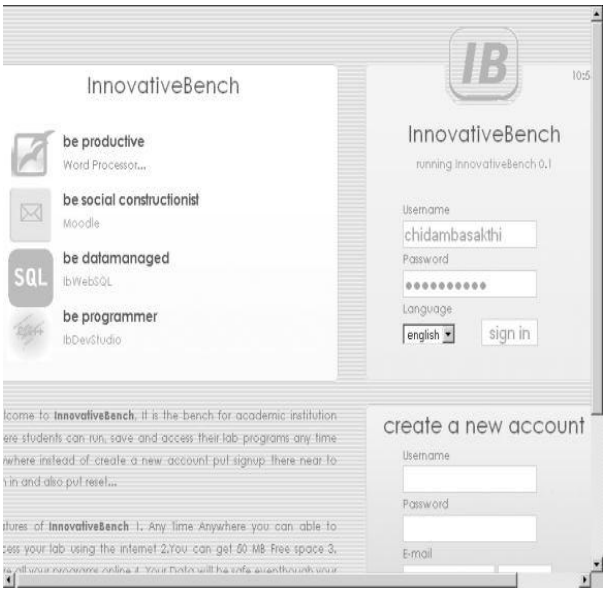
viii) IBCalendar: IBCalendar is used to set the timer and note down the event and to-do list.

V. RESULTS

To access the virtual lab, we must login to the index and the details which needed here is so simple and using that the whole details of a user is to be maintained.

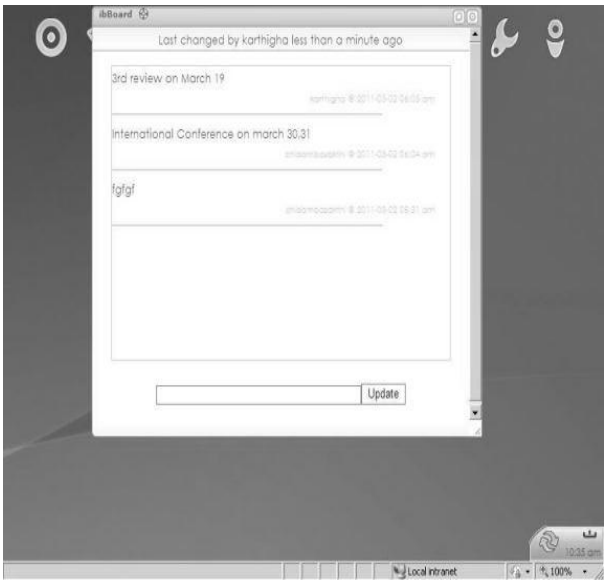
The user can able to access the various type of application with a single id and the new user can able to workout easily with the guidance screen which is in the flash format.

Figure-2: Login page of Virtual lab (Innovative Bench)



Sources: Authors Compilation.

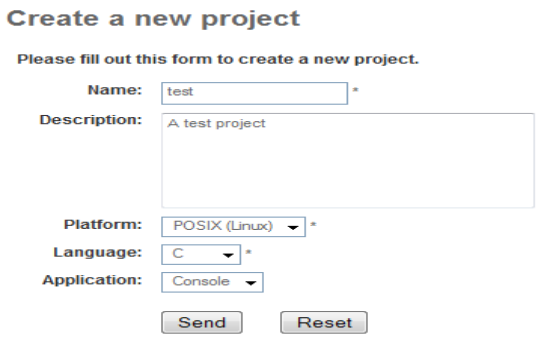
Figure-3: Web Desktop with IBBoard Application



Sources: Authors Compilation.

IBBoard will run and check if the board has been changed since the last time it recorded you displaying it. If no changes have been made, IBBoard will terminate. If the board has been updated it will be displayed as normal.

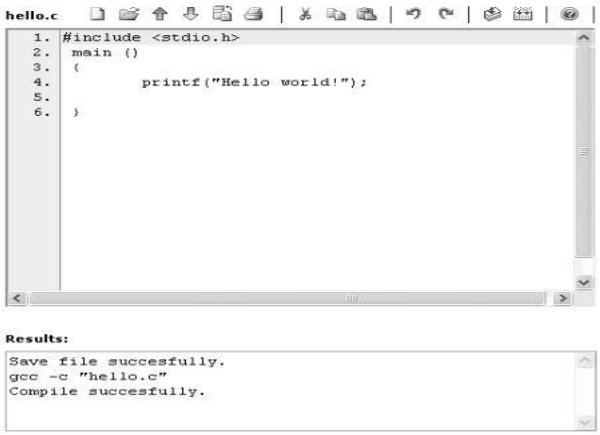
Figure-4: IBDeveloper Request Page



Sources: Authors Compilation.

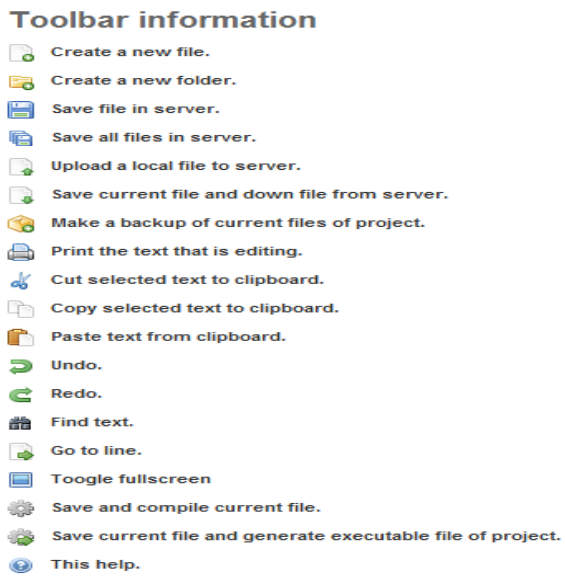
To design programs without greater complexity, using only a Web navigator and without worrying about file and project (creation, modification, compilation, etc.)

Figure-5: IBDeveloper Editor Page



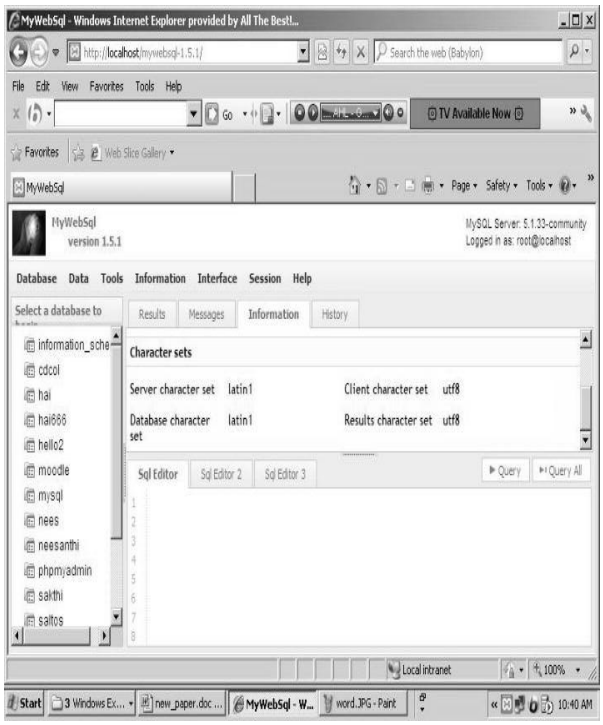
Sources: Authors Compilation.

Figure-6: IBDeveloper Toolbox



Sources: Authors Compilation.

Figure-7: IBWebsql Layout



Sources: Authors Compilation.

Maintaining and configuring any kind of database is possible using this websql and we can import and export any kind of database. This will able to avoid the configuration headaches and easy to create tables and all.

Figure-8: IBWebsql Database Creation Method

Sources: Authors Compilation.

Figure-9: IBEdit with Word & Table

Sources: Authors Compilation.

The Editor is very much essential for the office purpose and this ibedit have all the premises that the ms-word and open office have. We can able to export and import the local file to ibenvironment.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we present the ideas of the virtual lab and the implementation methodology of the virtual lab using the web based operating system. And introduce some of the engines that act as the kernel structure to perform the application over it.

The only problem with this virtual lab is the bandwidth and hopes the 3G and 4G introduction equate that pitfall. Due to the analysis, we inferred that hardware also plays a main role for the demonstration cost. If we go with the hardware less infrastructure, we will achieve the economical lab setup and that may be the future work that we proposed.

VII. REFERENCES

[1] Vishnu, S., Pendyala; Synopsys, Simon; S., and Y. Shim, (2009). "The Web as the Ubiquitous Computer, Web Technologies", Journal of IEEE Computer Society, September, pp. 90-92.

[2] Geng, Lin, (2009). "IBM Alliance at Cisco Systems, David Fu, Microsoft", Cloud Computing IT as a Service, March/April 2009, IEEE Journal of Pervasive Computing, pp. 10-14.

[3] George, Lawton, (2008). "Moving the OS to the Web", Journal of IEEE Computer Society, March, pp. 16-19.

[4] Marianne, C., Murphy, and Martin, Mcclelland, "Computer Lab to Go: A "Cloud" Computing Implementation", Proc ISE Con, 2008.

[5] Anisett, M.; Bellandi, V.; Colombo, A.; Cremonini, M.; Damaini, E., and Frati, F., et al. (2007). "Learning Computer Networking on Open Paravitual Laboratories", IEEE Transactions on Education, 50(4), pp. 302- 311.

[6] Averitt, S.; Vouk, M.; Schaffer, H.; Sills, E.; and Howell, G., (2004). "Support for Distributed Scalable High Performance and Grid-based Computing and Problem Solving with Emphasis in Science and Engineering", North Carolina State University: UNCG.

[7] Gaspar, A.; Langevin, S.; Armitage, W.; Sekar, R., and Daniels, T., (2008). "The Role of Virtualization in Computing Education", ACM SIGCSE Bulletin, 40(1).

[8] Seay, C., and Tucker, G., (2007). "The Virtual Computing Initiative at a Small Public University", North Carolina Central University.

[9] T., Freeman et al., (2006). "Division of Labor: Tools for Growth and Scalability of the Grids", Proc. 4th Int'l Conf. Service-Oriented Computing, Springer, pp. 40–51.

[10] M., Bailey, et al., (2007). "The Data Center of the Future", Int'l Data Group, Doc. No. 06C4799, April.

[11] G., B. Harwood, and R., H. Hermanson, (1976). "Lease-or-Buy Decisions", J. Accountancy, Volume: 142, Number: 3, 1976, p. 83-87.

[12] www.blog.eyeos.org

[13] www.skytap.com

[14] www.vcl.ncsu.edu

[15] www.vcl.unc.edu

[16] www.icloud.com

[17] www.moodle.org

*****

AN ANALYSIS OF VLSI DESIGN BASED ON EFFICIENT MULTIPLIER ARCHITECTURE

M. Srinivasan[55] R. S. Kamalakannan[56]

ABSTRACT

A multiplier is designed with high-speed, low power and customary in layouts that have generous research interest. This hybrid multiplier is designed by combination of multiplier and Spurious Power Suppression Technique. SPST technique reduced the dynamic power dissipation. Basically the survey is done on certain performance parameters such as Area, Speed, and Power consumption. In multiplier-and-accumulator (MAC) the Multipliers are to be considered as an important component in digital signal processing Applications like filters i.e. IIR. Therefore, the low power multiplier plays a vital role in the design and implementation of efficient power-aware devices. In this paper we have analyzed various multiplier architectures based on their working principles, Performance, area, and power consumption.

KEYWORDS

Modified Booth Encoding, Partial Product, Carry Save Multiplier, SPST Adder, Digital Signal Processing etc.

I. INTRODUCTION

In digital computer especially in signal processing system such as graphics and computation the multiplication is present. In VLSI implementation low power concept is necessary to meet Moore's law and to produce consumer electronics with more back up and less weight. It requires more hardware develop process independent chip design tools are growing because of the continuous development of VLSI technology [1], [2], [3]. At the early stages the hardware description language is good for the integrated circuit design. Multipliers are the basic component in single chip digital information process [5], [6]. Many scientists have tried and trying to design multiplier which offers high speed, low power consumption.

The use of very High speed hardware description language (VHDL) and Verilog. The area and speed are two important conflicting constrains. So Performance is always base on the area. In larger area the speed result is improved. The number of gates per chip area increasing, but the gate switching energy does not decreases at the same rate. So the power dissipation rises and removal of heat becomes difficult and expensive. There are different multiplier structures which can be classified as serial multiplying parallel multiplier, array multiplier, tree multiplier, it is categorized based on their architecture, application, and the way of producing partial products and summing up of partial products.

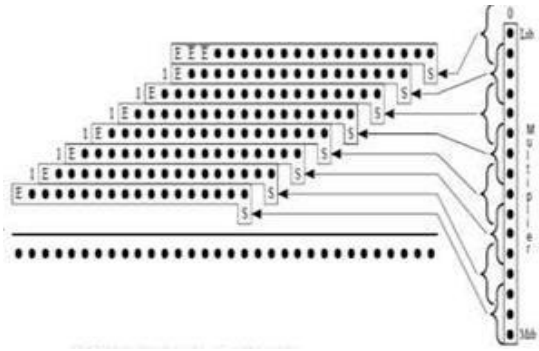
[55] Assistant Professor, Shree Venkateshwara Hi-Tech Engineering College, Tamil Nadu, India, kmsrinivasan@gmail.com

[56] Assistant Professor, Shree Venkateshwara Hi-Tech Engineering College, Tamil Nadu, India, rskamal09@gmail.com

II. MODIFIED BOOTH ENCODER

Booth Encoding is one of the techniques for to reducing the partial products required to producing the multiplication result. The main purpose of the algorithm is to achieve the high speed multiplication, algorithm using parallel counter as like modified booth algorithm has been proposed. Modified booth multipliers allow for smaller, faster multiplication circuits by using booth recoding technique [18] the number of partial products is reduced to half. Reduction in the number of partial products depends up on how May bits are recoded and on grouping of bits.

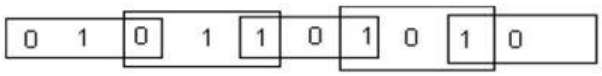
Figure-1: Modified Booth Encoder



Sources: Authors Compilation.

Grouping considered each 3 bits starts from LSB and first considered only two bits, next considered 3 bits but one bit will be overlapped on the previous group,

Figure-2: Grouping of Bits from Multiplier Term



Sources: Authors Compilation.

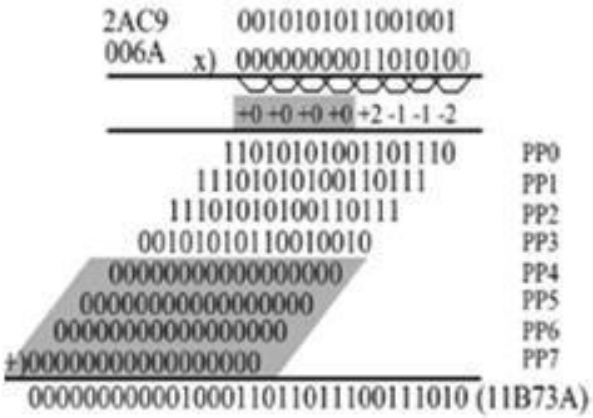
Table-1: Recoding Table

The advantages of the modified booth multiplier is reducing the number of partial products in to half of the multiplier term size by grouping, complexity of the circuit to produce partial product is the main disadvantages of this multiplier.

Block	Re - coded digit	Operation
000	0	0
001	+1	+1
010	+1	+1
011	+2	+2
100	-2	-2
101	-1	-1
110	-1	-1
111	0	0

Sources: Authors Compilation.

Figure-3: Grouping of Bits from Multiplier Term



Sources: Authors Compilation.

III. SPURIOUS POWER SUPPRESSION TECHNIQUES AND REGISTER

Spurious transitions (also called glitches) in combinational CMOS logic are a well known source of unnecessary power dissipation. Reducing glitch power is a highly desirable target because in the vast majority of digital CMOS circuits, only one signal transition per clock cycle are functionally meaningful.

Unfortunately, glitch power is heavily dependent on the low-level implementation details, namely, gate propagation delays and input transitions misalignments. The procedure for glitch minimization is based on a well known idea. Glitches are eliminated by International Journal of VLSI design & Communication Systems (VLSICS) Vol.3, No.3, June 2012 112 adding some redundant logic that prevents spurious transitions. This can be done by inserting latches in a gate-level net list.

Figure-3 shows a 16-bit adder / subtractor based on the proposed SPST. In this, the 16-bit adder / subtractor are divided into MSP (Most Significant Part) and LSP (Least Significant Part) between the 8th and 9th bits. The MSP of the original adder is modified to include detection logic circuits, data controlling circuits, sign extension circuits, logic for calculating carry-in and carry-out signals.

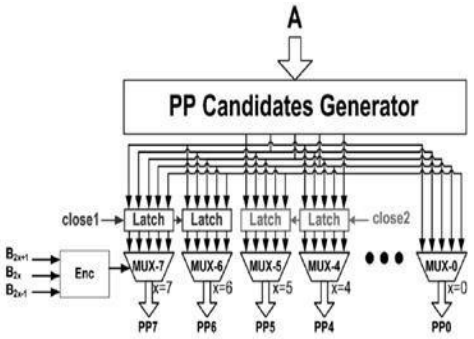
Simple logic gates are used to implement the latches and the sign extension circuits in order to reduce the additional overhead as for as possible. Low power adder / subtractor consist of: i) latch, ii) Detection logic, and iii) sign extension logic.

Applying SPST toModified Booth Encoder

The SPST equipped modified Booth encoder, which is controlled by a detection unit. One of two operands as input to the detection unit, which decide whether booth encoder calculates redundant computations.

As shown in Figure-4, the latches can, respectively, freeze the inputs of MUX-4 to MUX-7 or only those of MUX-6 to MUX-7 when PP4 to PP7 or PP6 to PP7 are zero, to reduce the transition power dissipation.

Figure-4: SPST Equipped Modified Booth Encoder.



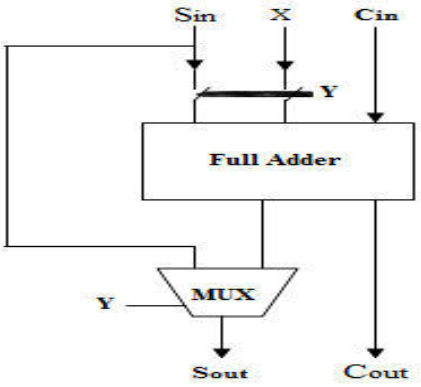
Sources: Authors Compilation.

IV. MIXED STYLE MULTIPLIER

Low power design is satisfy Moore's law and to produce low power consumption and high batter pack-up. In this modern world the saving power is necessary. In the technology above 0.1m dynamic power is the dominant property. Leakage power is more important in the smaller technologies charging the load capacitance in a circuit it will occur due to dynamic power dissipation.

By adding transient power consumption [PT] and capacitive load power consumption [PL] to calculate the dynamic power consumption of CMOS Ic for the past few years to minimize the dynamic power dissipation in arithmetic circuits, [19] [20] to apply different optimizations.

Figure-5: The FAB Cell

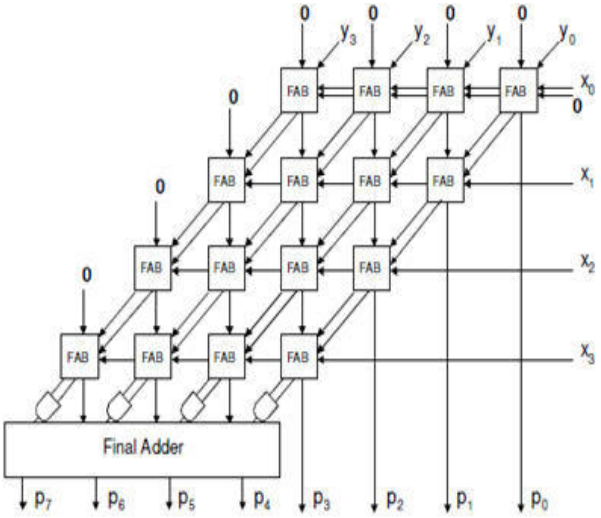


Sources: Authors Compilation.

Two kinds of multiplier are available in the mixed style multiplier, there are array part and tree part multiplier. When we choose the array part carry save array multiplier will provide better result when it's combined with the bypassing technique. By using bypassing technique the design of low power combinational circuit is introduced but their function is not required. This bypassing is using low delay and area overhead component.

This technique avoide switching activity in circuit and saving more dynamic power. This kind of multiplier has linear delay circuit [21]. This bypassing technique is applied into the array part of the multiplier.

Figure-6: The Carry-save Array Multiplier with Bypass



Sources: Authors Compilation.

X= (X_{n-1}, X_1, X_0) and Y= (Y_{m-1}, Y1, Y0) are the functionality of carry save array multiplier are fed in to FAB cell. When y=0 the multiplexer propagate the Sin and Sout. At this time the transmission gate in FAB cell locked. Input of full adder to prevent certain operation. When y=1 the full adder work and produce Sout. Another one is tree part multiplier. Logarithmic circuit delay and produce results in faster way is the advantage of the tree part multiplier.

The SPST adder and carry save array multiplier can be combined in mixed multiplier architecture. The power consumption and dynamic power saving this multiplier give good result, but time and area it is not sure give good result. It's irregular and complex interconnection it's hard to implement Wallace tree in FPGA. By using faster and efficient multiplier it give better result.

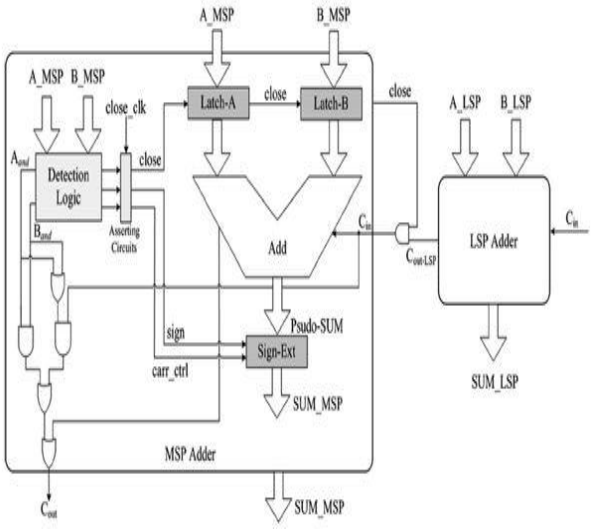
V. PROPOSED ARCHITECTURE

The Carry save array multiplier is a one type of dynamic power saving multiplier, here power saving is obtained by using the bypass technique across full adders. It can be used for only the low power consumption components such as transmission gate, multiplexer in the place of full adder.

So it reduces the consumption power. It avoids the unwanted switching activity which in turn reduces the dynamic power dissipation. We can expect large amount power savings in carry save array multiplier. On the other hand, in all cases the proposed bypass architecture offers power savings ranging from 20% up to almost 60%.

A new architecture can be proposed by combining bypass carry save array multiplier with SPST adder it may produce results in a much faster way and reduce the dynamic power dissipation. Since Carry save array multiplier with SPST is considered to have best result in Area-Delay2 product (AD2) and Delay Product (DP) among some of the multiplier. Considering Wallace tree multiplier in the place of SPST adder may give good and fast results and will reduce power dissipation.

Figure-7: Low-Power Adder Design Using SPST



Sources: Authors Compilation.

VI. RECOMMENDATION & LIMITATIONS

The proposed multiplier is designed for all parameters like as time and area with power should be chosen. Combination of these multiplier changes gives better result comparing to individual multiplier. Based on the design consideration each multiplier has it's on own characteristics. A proposed architecture is based on hybrid type multiplier. The hybrid type multiplier is designed by combining the SPST adder with multiplier, it will not provide better performance as like as the individual multiplier.

VII. CONCLUSIONS & FUTURE WORK

Considering the all other multiplier hybrid type of multiplier gives better result. The hybrid multiplier is designed by combination of multiplier with SPST. The SPST adder avoids the unwanted glitches and thus minimizes the switching power dissipation. The main advantage of this multiplier is dynamic power saving in low power VLSI design.

The implemented radix-2 modified Booth Algorithm MAC with SPST gives a factor of 5 less delay and 7% less power consumption as compared to array MAC. Radix-2 modified booth algorithm reduces the number of partial products to half by grouping of bits from the multiplier term, which improves the speed. When we considering the ASCI implementation we can be further extend this work with analysis of power and area.

VIII. REFERENCES

[1] S., M. Aziz, and Iftekhar, Ahmed, (1998). "Easily Testable Array Multiplier Design Using VHDL", Malaysian Journal of Computer Science, December, Volume: 11, Number: 2, pp. 1-7.

[2] D., D. Gazski; N., D. Dutt; C., H. Wu, and Y., L. Lin, (1991). "High-Level Synthesis, Introduction to Chip and System Design", Kluwer Academic Publishers.

[3] Z., Navabi, (1993). "VHDL Analysis and Modeling of Digital Systems", New York, McGraw-Hill, Inc.

[4] J., R., Armstrong, (1989). "Chip Level Modeling with VHDL", Englewood Cliffs, New Jersey: Prentice-Hall International.

[5] Frank, P. J.; M., Welton, and Antoine, Delaruelle, (1985). "A 2-m m CMOS 10-MHz Micro Programmable Signal Processing Core With an On-Chip Multiport Memory Bank", IEEE Journal of Solid-State Circuits, Volume: SC-20, Number: 3, June, pp.754-760.

[6] K., Takeda; F., Ishino, and Y., Ito et al., (1985). "A Single-Chip 80-bit Floating Point Processor", IEEE Journal of Solid-State Circuits, Volume: SC-20, Number: 5, October, pp. 986-991.

[7] N., F., Benschop, (1983). "Layout Compilers for Variable Array Multipliers", in Proc. of Custom Integrated Circuits Conf., May, pp. 336- 339.

[8] K., C. Chu, and R., Sharma, (1984). "A Technology Independent MOS Multiplier Generator", in 21st Design Automation Conf., pp. 90-97.

[9] S., Shah; A., J. Aj-Khabb, and D., AI-Khabb, (2000). "Comparison of 32-bit Multipliers for various Performance Measures", 12th International Conference on Microelectronics Tehran, October 31 - November 2.

[10] C., S. Wallace, (1964). "A Suggestion for a Fast Multiplier", IEEE Trans. Elechon. Con@., Volume: EC-13, pp. 14-17, February.

[11] L., Dadda, (1965). "Some Schemes for Parallel Multipliers", in Proc. Alta Frequenza, March, Volume: 19, pp. 349-356.

[12] Mahmoud, A. Al-Qutayri; Hassan, R. Barada, and Ahmed Al-Kindi, "Comparison of Multipliers Architectures through Emulation and Handle-C FPGA Implementation", Etisalat University College, Sharjah, UAE.

[13] P., Meier; R., A., Rutenbar; and L., R. Carley, (1996). "Exploring Multiplier Architecture for Low Power", in Custom Integrated Circuits Conference IEEE, pp. 513-516.

[14] G., Economakos, and K., Anagnostopoulos, (2006). "Bit Level Architectural Exploration Technique for the Design of Low Power Multipliers", in International Symposium on Circuits and Systems, IEEE.

[15] T., Sakuta; W., Lee, and P., Balsara, (1995). "Delay Balanced Multipliers for Low Power / Low Voltage DSP core", in Symposium on Low Power Electronics. IEEE, pp. 36-37.

[16] J., Rabaey; A., Chandrakasan, and B., Nikolic, (2003). "Digital Integrated Circuits: A Design Perspective", 2nd Ed., Prentice Hall.

[17] A., Bellaouar, and M., Elmasry, (1995). "Low Power Digital VLSI Design: Circuits and Systems", Kluwer Academic Publishers.

[18] C., N. Marimuthu1, and P., Thangaraj, (2008). "Low Power High Performance Multiplier", ICGST-PDCS, Volume: 8, Issue: 1, December.

[19] M., Karlsson, (2000). "A Generalized Carry-Save Adder Array for Digital Signal Processing", in 4th Nordic Signal Processing Symposium, IEEE, pp. 287–290.

[20] S., M. Lee; J. H. Chung; H., S. Yoon, and M., M. O. Lee, (2000). "High Speed and Ultra Low Power 16x16 MAC Design Using tg Techniques for Web based Multimedia System", in Asia South Pacific Design Automation Conference. ACM/IEEE, pp. 17–18.

[21] N. Weste, and D., Harris, CMOS VLSI Design: A Circuits and Systems Perspective, 3rd Edition, Addison-Wesley.

[22] Dimitris, Bekiaris; George, Economakos, and Kiamal Pekmestzi, (2010), "A Mixed Style Multiplier Architecture for Low Dynamic and Leakage Power Dissipation", National Technical University of Athens. IEEE.

*****

OPTIMIZED INTENTION FOR CONTINUOUS QUERIES IN THE ACTIVE DATA AGGREGATION NETWORK WITH COST MODEL

Indumathi R.[57] N. Mookhambika[58]

ABSTRACT

Within an RDBMS streams of changes to the data and reporting when the result of a query defined over the data changes. These queries are referred to as the continuous queries since they continually produce results whenever new data arrives or existing data changes. To make online decision we require monitoring the continuous queries.

Here the aim is to introduce the low-cost, scalable technique to answer continuous aggregation queries using a network of aggregators of dynamic data items. There is significant work in systems that can efficiently deliver the relevant updates automatically and also to provide for getting the optimal set of sub queries with their incoherency bounds which satisfies client query's coherency requirement with least number of refresh messages sent from aggregators to the client. For optimal query execution divide the query into sub-queries and evaluate each sub-query at a judiciously chosen data aggregator.

The main purpose is to response the client with the least number of tasks with the help of random query selection. Random query selection means for the user submitted query the relevant queries, sub queries are generated. A query cost model which can be used to estimate the number of messages required to satisfy the client specified incoherency bound.

KEYWORDS

Query, RDBMS, Data, Aggregation, Client etc.

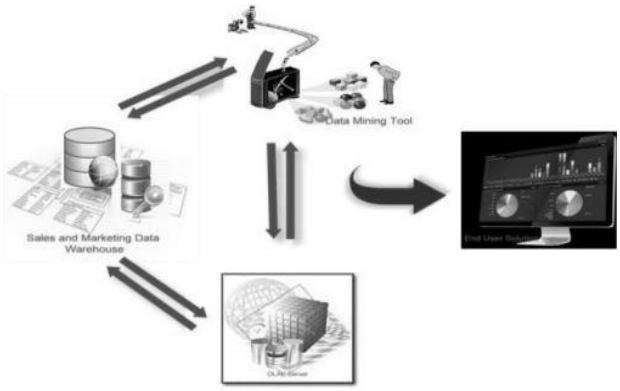
INTRODUCTION

Data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. While large-scale information technology has been evolving separate transaction and analytical systems, data mining provides the link between the two.

Data mining software analyzes relationships and patterns in stored transaction data based on open-ended user queries. Several types of analytical software are available: statistical, machine learning, and neural networks. Data mining consists of five major elements:

1) Extract, transform, and load transaction data onto the data warehouse system.
2) Store and manage the data in a multidimensional database system.
3) Provide data access to business analysts and information technology professionals.
4) Analyze the data by application software.
5) Present the data in a useful format, such as a graph or table.

Figure-1: Integrated Data Mining Architecture



Sources: Authors Compilation.

The Figure-1 represents a fundamental shift from conventional decision support systems. Rather than simply delivering data to the end user through query and reporting software, the Advanced Analysis Server applies users' business models directly to the warehouse and returns a proactive analysis of the most relevant information. These results enhance the metadata in the OLAP Server by providing a dynamic metadata layer that represents a distilled view of the data. Reporting, visualization, and other analysis tools can then be applied to plan future actions and confirm the impact of those plans.

The more powerful the data mining queries, the greater the utility of the information being gleaned from the data, and the greater the pressure to increase the amount of data being collected and maintained, which increases the pressure for faster, more powerful data mining queries. This increases pressure for larger, faster systems, which are more expensive. All database systems must be able to respond to requests for information from the user—i.e. process queries. Obtaining the desired information from a database system in a predictable and reliable fashion is the scientific art of *Query Processing*.

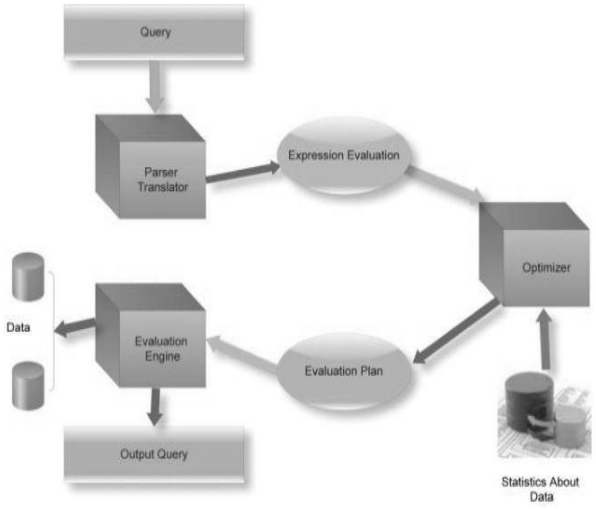
Getting these results back in timely manner deals with the technique of *Query Optimization*. A database query is the vehicle for instructing a DBMS to update or retrieve specific data to/from the physically stored medium. The actual updating and retrieval of data is performed through various "low-level" operations. There are three phases that a query passes through during the DBMS' processing of that query: Parsing and translation, Optimization and Evaluation.

Most queries submitted to a DBMS are in a high-level language such as SQL. During the parsing and translation stage, the human readable form of the query is translated into forms

[57] Student, Final Year, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, indu_be@yahoo.co.in
[58] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

usable by the DBMS. These can be in the forms of a relational algebra expression, query tree and query graph.

Figure-2: Query Processing



Sources: Authors Compilation.

RELATED WORKS

Data Stream Management Systems (DSMSs) were developed to be at the heart of every monitoring application. Monitoring applications typically register hundreds of Continuous Queries (CQs) in DSMSs in order to continuously process unbounded data streams to detect events of interest. DSMSs must be designed to efficiently handle unbounded streams with large volumes of data and large numbers of CQs, i.e., exhibit scalability [6]. This need for scalability means that the underlying processing techniques a DSMS adopts should be optimized for high throughput. These systems are built around three design principles that aid in the real-time. Querying of complex data sources: query interfaces tailored to the application's specific data types, optimized data collection processes, and allowing queries to provide feedback to the collection process.

Various mechanisms for efficiently maintaining incoherency bounded aggregation queries over continuously changing data items are proposed in the literature [5, 7, 9]. This paper work distinguishes itself by being sub-query based evaluation to minimize number of refreshes. In [7], authors propose using data filters at the sources; instead here assign incoherency bounds to sub-queries which reduce the number of refreshes for query evaluation. In proxy-based caching approaches store content at various locations outside the site infrastructure and can improve Web site performance by reducing content generation delays, firewall processing delays, and bandwidth requirements [2]. In cache at the page level, which does not guarantee that correct pages axe served and provides very limited reusability. The benefits of existing proxy-based and back end caching techniques, without their respective limitations.

In traditional database systems optimize for performance on one-shot query processing, emerging large-scale monitoring applications require continuous tracking of complex data analysis queries over collections of physically-distributed streams. In readily incorporates several complex analysis queries thus giving the first known low-overhead tracking solution for such queries in the distributed-streams model [8]. The optimized for tracking high-speed streams, and result in very low processing and communication costs, and significant savings over naive updating schemes.

The method of executing queries over dynamic data dissemination network is practical since it can be implemented using a mechanism similar to URL-rewriting [4] in CDNs. Just like in a CDN, the client sends its query to the central site. For getting appropriate aggregators (edge nodes) to answer the client query (web page), the central site has to first determine which data aggregators have the data items required for the client query. If the client query cannot be answered by a single data aggregator, the query is divided into sub-queries (fragments) and each sub-query is assigned to a single data aggregator.

PROPOSED SYSTEM

The Web suffers from performance and scalability issues. Content distribution networks (CDNs) solved the problem for static content using caches at the edge nodes of the networks. CDNs continue to evolve to serve more and more dynamic applications. A dynamically generated web page is usually assembled using a number of static or dynamically generated fragments. The static fragments are served from the local caches whereas dynamic fragments are created either by using the cached data or by fetching the data items from the origin data sources. One important question for satisfying client requests through a network of nodes is how to select the best node(s) to satisfy the request.

The content requested proximity to the client and load on the nodes are the parameters generally used to select the appropriate node. In dynamic CDNs, while selecting the node(s) to satisfy the client request, the central site (top-level CDN node) has to ensure that page / data served meets client's coherency requirements also. Such dynamic data dissemination networks can be used to disseminate data such as stock quotes, temperature data from sensors, traffic information, and network monitoring data. In this paper we propose a method to efficiently answer aggregation queries involving such data items. Incoherency of a data item at a given node is defined as the difference in value of the data item at the data source and the value at that node. Incoherency bound, a data aggregator gets data updates from the data source or some higher level data aggregator so that the data incoherency is not more than the data incoherency bound. In hierarchical data dissemination network a higher level aggregator guarantees a tighter incoherency bound compared to a lower level aggregator. Thus data refreshes are pushed from data sources to the clients through the network of aggregators.

The problem of selecting the optimal plan ensure that each data item for a client query is disseminated by one and only one data aggregator. Although a query can be divided in such a way that a single data item is served by multiple; but in doing so the same data item needs to be processed at multiple aggregators, increasing the unnecessary processing load. By dividing the client query into disjoint sub-queries it ensure that a data item update is processed only once for each query.

The query incoherency bound needs to be divided among sub query incoherency bounds such that, besides satisfying the client coherency requirements, the chosen is capable of

satisfying the allocated sub-query incoherency bound. Here it can prove that the number of refreshes depends on the division of the query incoherency bounds among sub-query incoherency bounds. The individual data items involved. The data dissemination cost is dependent on data dynamics and incoherency bound associated with the data. We model the data dynamics using a data synopsis model, and the effect of the incoherency bound using an incoherency bound model. These two models are combined to get the estimate of the data dissemination cost.

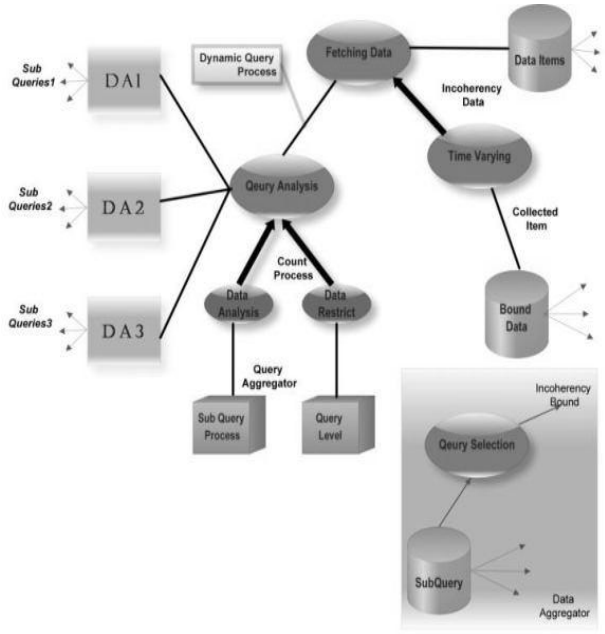
ARCHITECTURE OF SYSTEM

There are several types of web services available now a day, the services are provided by the different kind of service provider. The client can view or access the number of services from the service provider. And they can search as per their want to search.

For example they can login to the Gmail, twitter or some other else. And the user can give the numerous queries they want to be search. The service provider keeps the distinct queries submitted by the user. And also they keep the track of pages viewed. The service provider can update the query either by static or dynamic. And also they update the pages viewed by the users. For that submitted query, the answers cannot get by the single aggregator. The submitted query were split and searched by the number of aggregator.

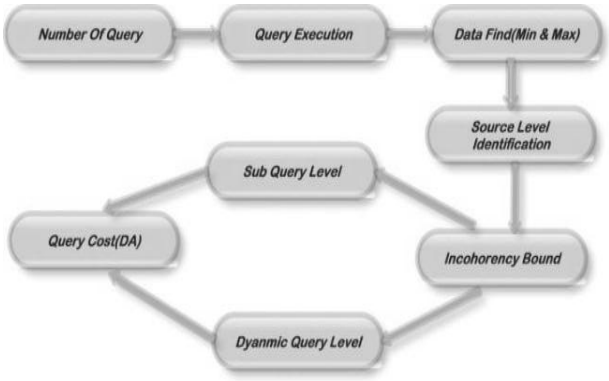
The results from the distinct aggregator were collected and response to the client. The above concept purpose is to response the client with the least number of tasks with the help of random query selection. Random query selection means for the user submitted query the relevant queries, sub queries are generated. In query cost execution, the costs of refreshes are updated.

Figure-3: Architectural View of System



Sources: Authors Compilation.

Figure 4: Evaluation of Query



Sources: Authors Compilation.

The algorithm is to execute the continuous query and to calculate the cost (i,e) to estimate the number of refresh messages sent from aggregators to the client. The outline of Greedy algorithm for delivering the queries is follow:

Algorithm GreadyBFS
Get input data
Complete a path tree for each origin query by doing the following.
Initialize the query table
Step 1: Create the list of tentative query labels
Update the node table
Step 2: Finalize the current query
Step 3: Find the next current query
if there is a valid current query then Go to Step 1
Otherwise, done
Output the completed query table

First, get a set of maximal sub-queries corresponding to all the data aggregator in the network. The maximal sub-query for a data aggregator is defined as the largest part of the query which can be disseminated by the DA. [3] For the given client query and the relation consisting of data aggregators, data-items, and the data incoherency bounds ,maximal sub queries can be obtained for each data aggregator by forming sub-query involving all data items in the intersection of query data items and those disseminated by the DA.

Algorithm for QueryExecutionWith CalculatingCost:
Begin
If have_input_data <> 1 Then
get_query
get_links
ReDim query _table(num_query, 4)
ReDim all_origins_ query _table(num_ query * num_zones, 5)
first_origin = 0
End If
have_input_data = 1
calc_costs
For node_count = 1 To num_ query
If nodes(node_count, 2) = "z" Then
origin_zone = node_count
initialize_node_table
debug_output
origin_done = "no"
Do Until origin_done = "yes"
find_neighboring_nodes
update_node_table
debug_output
finalize_current_node
debug_output

```
find_current_node
Loop
output_origin_node_table
if first_origin = 0 then
first_origin = 1
End if
End If
Next
End
```

For executing an incoherency bounded continuous query, a query plan is required which includes the set of sub-queries, their individual incoherency bounds and data aggregators which can execute these sub-queries [3]. The need to find optimal query execution plan which satisfies client coherency requirement with least number of refreshes. The mechanism to:

Task 1: Divide the aggregation query into sub-queries; and
Task 2: Allocate the query incoherency bound among them.
While satisfying the following conditions:

Condition1: Query incoherency bound is satisfied.
Condition2: The chosen DA should be able to provide all the data items appearing in the sub-query assigned to it.
Condition3: Data incoherency bounds at the chosen DA should be such that the sub-query incoherency bound can be satisfied at the chosen DA.
Objective: Number of refreshes should be minimized.

COST MODEL FOR AGGREGATION QUERIES

Consider a query over two data items P and Q with weight w_0 and w_1 and now is to estimate the dissemination cost: The query sumdiff will be:

$$R_{data} = w_p R_p + w_q R_q = w_p \sum | p_i - p_{i-1} | + w_q \sum | q_i - q_{i-1} |$$

The aggregator uses the information that client is, P & Q is an individual query value. User Compare Query Data Analysis[3],

$$R_{query} = \sum | w_p (p_i - p_{i-1}) + w_q (q_i - q_{i-1}) |$$

Thus to represent the relationship between *Rquery* and *sumdiff* values of the individual data items using a correlation measure associated with the pair of data items.

$$R^2_{query} \propto (w_p^2 R_p^2 + w_q^2 R_q^2 + 2\rho w_p R_p w_q R_q)$$

ρ =Correlation Measure.

$$\rho = \frac{\sum (p_i - p_{i-1})(q_i - q_{i-1})}{\sqrt{\sum (p_i - p_{i-1})^2} \sqrt{\sum (q_i - q_{i-1})^2}}$$

Compare Query between P & Q,

$$R^2_{query} = \frac{(w_p^2 R_p^2 + w_q^2 R_q^2 + 2\rho w_p R_p w_q R_q)}{(w_p^2 + w_q^2 + 2\rho w_p w_q)}$$

The value of the normalizing factor R_{query} :

$$1 / \sqrt{w_p^2 + w_p^2 + 2\rho w_p w_q} \ .$$

The value of the incoherency bound(C) has to be adjusted by the same factor. The query cost model we first summarize the model to estimate the number of refreshes required to disseminate a data item at certain incoherency bound.The number of data refreshes is inversely proportional to the square of the incoherency bound ($1/C2$).The data dynamics was modeled as a random walk process.

$$Data\ dissemination\ cost \ \propto \ 1/C^2$$

The number of update messages for a data item is likely to be higher if the data item changes more in a given time window. The estimated dissemination cost for data item *S*, disseminated with an incoherency bound *C*, is proportional to *Rs/C2*. Using this result the query cost model is developed.

SUMMARY OF PERFORMANCE RESULTS

Following features of the query planning algorithm improve performance:

- Dividing the query into sub-queries and executing them at specifically chosen data aggregators.
- Deciding query plan using data *sumdiff* based mechanism specifically by maximizing sub-query gains.
- Including more dynamic data as part of a larger sub-query.

CONCLUSIONS

This paper presents a cost based approach to minimize the number of refreshes required to execute an incoherency bounded continuous query. For optimal execution it is to divide the query into sub-queries and evaluate each sub-query at a chosen aggregator. In this method, the query execution can be implemented using schemes similar to that used in CDNs. The query cost model can also be used for other purposes such as load balancing various aggregators, optimal query execution plan at an aggregator node, etc.

REFERENCES

[1] Mohamed, F. Mokbel; Xiaopeng, Xiong, and Walid, G. Aref, "SINA: Scalable Incremental Processing of Continuous Queries in Spatiotemporal Databases", Department of Computer Sciences, Purdue University, West Lafayette, IN 47907-1398.

[2] D., Vander Meer; A., Datta; K., Dutta; H., Thomas, and K., Ramamritham, (2004). "Proxy-Based Acceleration of Dynamically Generated Content on the World Wide Web", ACM Transactions on Database Systems (TODS), June, Volume: 29.

[3] Rajeev, Gupta, and Krithi, Ramamrithram, "Query Planning for Continuous Aggregation Queries over a Network of Data Aggregators", IEEE Transaction on Knowledge and Data Mining, Volume: 24, Issue: 6

[4] S., Rangarajan; S., Mukerjee, and P., Rodriguez, (2003). "User Specific Request Redirection in a Content Delivery Network", 8th Intl. Workshop on Web Content Caching and Distribution (IWCW).

[5] R., Gupta; A., Puri, and K., Ramamritham. (2005). "Executing Incoherency Bounded Continuous Queries at Web Data Aggregators", www.google.com

[6] Mohamed, A. Sharaf; Alexandros, Labrinnidis, and Panos, k. Chrysanthis, "Scheduling Continuous Queries in Data Stream Management Systems" in VLDB Endowment.

[7]. C., Olston; J., Jiang, and J., Widom, (2003). "Adaptive Filter for Continuous Queries over Distributed Data Streams", SIGMOD.

[8] G., Cormode, and M., Garofalakis, (2005). "Sketching Streams through the Net: Distributed Approximate Query Tracking", VLDB.

[9] S., Zhu, and C., Ravishankar, (2004). "Stochastic Consistency and Scalable Pull-Based Caching for Erratic Data Sources". VLDB.

*****

CONDITION MONITORING OF INDUCTION MOTOR BALL BEARING USING MONITORING TECHNIQUES

B. Hulugappaa[59] Tajmul Pasha[60] K. M. Ramakrishna[61]

ABSTRACT

Rolling element bearings are critical components in induction motors and monitoring their condition is important to avoid failures. Several condition monitoring techniques for the bearings are available. Out of these, stator current monitoring is a relatively new technique. Vibration, stator current, acoustic emission and shock pulse methods (SPMs) and FEM, surface analysis, for the detection of a defect in the Inner race of induction motor ball bearing have been compared. The measurements were performed at different loads and different speeds. The defect in the bearing could be detected by all the methods.

KEYWORDS

Rolling, Bearings, Motors, Current, Loads etc.

I. INTRODUCTION

Induction motors are widely used in industry and are considered as critical components for electric utilities and process industries. In the case of induction motors, rolling element bearings are overwhelmingly used to provide rotor support. Although induction motors are reliable, they are subjected to some modes of failures. Based on studies according to Motor Reliability Working Group (MRWG) and the investigation carried out by Electric Power Research Institute (EPRI), a common mode of failure of an induction motor is the bearing failure followed by stator winding and rotor bar failures. The bearing failure increases the rotational friction of the rotor. Even under normal operating conditions of balanced load and good alignment, fatigue failure begins with small fissures, located below the surfaces of the raceway and rolling elements, which gradually propagate to the surface generating detectable vibrations and increasing noise levels. Continued stressing caused fragments of the material to break loose producing localized fatigue phenomena known as flaking or spalling. Electric pitting or cracks due to excessive shock loading are also among the different types of bearing damage described in the literature [1,2]

The widespread application of rolling element bearings in both industry and commercial life require advanced technologies to efficiently and effectively monitor their health status. There are many condition monitoring methods used for detection and diagnosis of rolling element bearing defects: vibration measurements, temperature measurement, shock pulse method (SPM) and acoustic emission (AE). Among these vibration measurements are most widely used. Even though the emphasis is on vibration measurement methods, stator current harmonics measurement is appearing as an alternative to the vibration measurement methods. In fact, large electrical machine systems are often equipped with mechanical sensors, which are primarily vibration sensors such as proximity probes. However, these are delicate and expensive. Various researchers have suggested that stator current monitoring can provide the same indications without requiring access to the motor. This technique utilizes results of spectral analysis of the stator current or supply current of an induction motor for the diagnosis [1].

A detailed review of different vibration and acoustic methods, such as vibration measurements in time and frequency domains, sound measurement, the SPM and the AE technique for condition monitoring of rolling bearings have been presented by Tandon and Choudhury [3].

Each bearing has a characteristic rotational frequency. With a defect on particular bearing element, an increase in vibration energy at this element's rotational defect frequency may occur. These characteristic defect frequencies can be calculated from kinematics considerations, i.e. geometry of the bearing and its rotational speed. For normal speeds, the bearing characteristic defect frequencies lie in the low-frequency range and are usually less than 500 Hz.The relationship of the bearing vibration to the stator current spectra can be determined by remembering that any air gap eccentricity produces anomalies in the air gap flux density. Since ball bearings support the rotor, any bearing defect will produce a radial motion between the rotor and stator of the machine.

Riley et al. [4] presented a method for sensor less on-line vibration monitoring of induction machines. This method assumed a linear relationship between the current harmonics and vibration level. Da-Ming Yang and James Penman [5] addressed the use of stator current and vibration monitoring to diagnose bearing condition. In their study it has been reported that monitoring of stator current provides an alternative method for diagnosing bearing condition that is generally less intrusive, simpler and successful detection of motor bearing condition is possible using line current sensing.

AE is the phenomena of transient elastic wave generation due to a rapid release of strain energy caused by structural alteration in a solid material under mechanical or thermal stresses. Generation and propagation of cracks are among the primary sources of AE in metals. AE transducers are designed to detect the very high frequency (450 kHz) stress waves that are generated when cracks extend under load.

The most commonly measured AE parameters are peak amplitude, counts and events of the signal. Counts involve counting the number of times the amplitude exceeds a preset voltage level in a given time and gives a simple number characteristic of the signal. An event consists of a group of counts and signifies a transient wave. One of the advantages of AE monitoring is that it can detect the growth of subsurface cracks. Hence, it is an important tool for condition monitoring. It has been shown that AE parameters can detect defects before they appear in the vibration acceleration [6–8].

Tandon and Nakra [9] demonstrated the usefulness of some AE parameters, such as peak amplitude and count, for detection of

[59] Assistant Professor, National Institute of Technology, Karnataka, India, hulugappab@indiatimes.com
[60] Associate Professor, National Institute of Technology Karnataka, India, bhniemech@gmail.com
[61]Professor, P.E.S. College of Engineering, Karnataka, India, rkpesce@gmail.com

defects in radially loaded ball bearings at low and normal speeds. Tan [10] has presented the application of AE for the detection of bearing failures. He has suggested that the measurement of area under the amplitude curve is preferred method for detection of defects in rolling element bearings.

The shock pulses caused by the impact in bearings initiate damped oscillations in transducer at its resonant frequency. Measurement of the maximum value of the damped transient gives an indication of the condition of rolling bearings. Low-frequency vibrations in the machine, generated by sources other than rolling bearings, are electronically filtered out. The maxim normalized shock value is a measure of the bearing condition. Shock pulse meters are simple to use so that semiskilled personnel can operate them. They give a single value indicating the condition of the bearing straightaway, without the need for elaborate data interpretation as required in some other methods. The principle is based on the fact that structural resonances are excited in the high-frequency zone due to impulsive loading caused, for example, from spalling of the races or rolling elements and can be detected by a transducer whose resonance frequency is tuned to it. The SPM, which works on this principle, uses piezoelectric transducer having a resonant frequency based at 32 kHz [11,12]. In industries SPM has gained wide acceptance for detecting the rolling element bearing defects.

The literature indicates that even though the emphasis is on vibration measurement methods for the detection of defects in rolling element bearings in induction motor, stator current harmonics measurement can be an effective alternative to the vibration measurement because of its faster diagnosis. Very few studies have been carried out on stator current monitoring of an induction motor for the detection of defects in these rolling bearing along with vibration monitoring and other condition monitoring methods.

Hence, there is a need for a comprehensive study of induction motor rolling element-bearing faults detection using stator current harmonics measurement in combination with vibration, AE and SPM condition monitoring techniques. So the present work was undertaken for the detection and diagnosis of induction motor rolling element-bearing faults have been carried out using vibration monitoring, AE and shock pulse along with stator current harmonics measurements. Experimental investigation has been carried out to study the changes in these parameters for bearings in good condition and with simulated defects in the outer race of the bearings of an induction motor.

2. EXPERIMENTAL SET-UP AND MEASUREMENTS

2.1. Test Rig

The test rig used in the present research work consists of a 1.1kW/1440 rpm single-phase induction motor driving the V-belt drive. The test bearing is the drive end bearing of induction motor, housed within the drive end cover plate of motor to support rotor. The rotor shaft of motor is extended to the left of drive end cover plate of motor for providing transmission of power through pulley drive.

The test bearing can be radial loaded with V-belt drive with the loading system. Vibration isolation rubber sheets were provided under the motor and its supporting legs to reduce the vibration transmission from ground to the test bearing. The drive to the induction motor is provided by a.c. power supply. The transducers for measurement of vibration, AE and SPM have been mounted in the zone of maximum load on to the drive end bearing of induction motor.

The Schematic diagram of the experimental set-up is shown in Fig. 1. The test rig has been designed to withstand a maximum load of 27 kg based on the rated power of induction motor for 1.5 HP. The test bearing of induction motor in this study is normal clearance, deep groove ball bearings SKF 6205. Equal amount of SKF grease (4 g) was applied to the bearing, which was immersed in kerosene and finally cleaned with acetone to remove preserving oil.

Figure-1: Schematic Diagroam of Test Ring

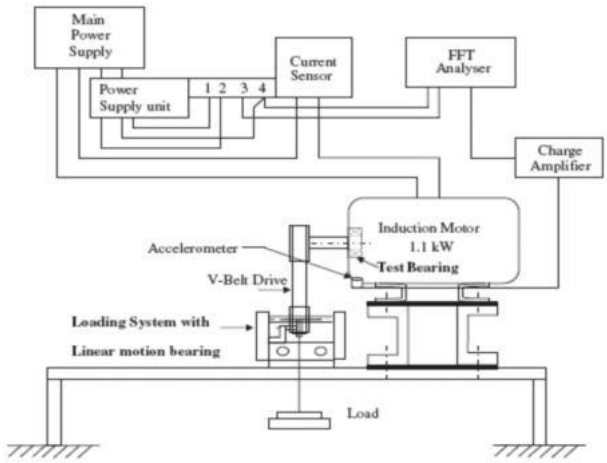


Fig. 1. Schematic diagram of test rig.

Sources: Authors Compilation.

2.2. Instrumentation

Vibrations were measured with the help of a piezoelectric accelerometer Bruel and Kjaer (B&K) type 4366having un damped natural frequency of 39 kHz. The output of the accelerometer was fed to the B&K charge amplifier 2635 connected to Ono Sokki CF 3200 portable fast Fourier transform analyzer. The schematic diagram of the current sensor (working on Hall Effect) in series with motor supply line is shown in Figure-1.

The Hall element located in the air gap of the magnetic circuit converts the magnetic field generated by the primary current into a proportional Hall voltage. The magnetic field produced by the primary current generates a highly linear magnetic flux in the air gap of the magnetic circuit, which in turn induces a proportional Hall voltage in the Hall element.

The voltage is then electronically amplified resulting in an output voltage that is highly proportional to the primary current up to the final value of the measuring range. The current sensor is supplied with 715V from the power supply unit for 15min before taking the measurements. The current flowing in single-phase induction motor was sensed by a current sensor of type LEM-HY 25P by connecting it in series with the supply line to the motor.

This current sensor has dynamic performance accuracy of linearity better than 70.2%, response time better than 1 ms and

nominal analog output current 25 m A. The output from the current sensor is fed to the FFT analyzer with frequency span of500 Hz and 20 kHz for frequency spectrum analysis and for overall values, respectively.AE measurements were performed by using Acoustic Emission Technology Corporation (AET), model AET AC 375 L transducer which are of same frequency of 375 kHz, a preamplifier with 60 dB gain (AET 160B) and a filter (AET FL 25) with a pass band of 250–500 kHz. The AE transducer was mounted on the test bearing housing with grease as couplant and with the help of cloth adhesive tape. The preamplifier is provided with 712V DC supply. The output from the preamplifier was fed to Tektronix TDS 210 digital real time oscilloscope, which can also give frequency spectrum.

The Shock Pulse Tester T2000 by SPM Instrument AB, Sweden, along with its hand held transducer type SPM 10777 was used for shock pulse measurements. The hand held probe was pressed straight at the zone of maximum load on bearing housing to get maximum normalized shock pulse value, dBM.

2.3. Measurement Conditions

The measurements were carried out from no load to full load (27 kg) for the induction motor bearing with an increment of 5 kg. The motor was run at constant speed of 1440 rpm. Three healthy bearings were used to check the repeatability of the measurements. Inner race defect was simulated by a circular hole of diameter varying from 250 to 1500 mm in outer race of the same bearing (in steps of 250 mm successively after each measurement) by spark erosion technique.

3. RESULTS AND DISCUSSIONS

3.1. Vibration Velocity

Figure-2 shows that the overall amplitudes of vibration velocity of three healthy bearings are very much close to each other and their average overall level is also shown. Figure-3 shows that overall velocity values follow the same trend as that of the good bearing with increase in load. The overall velocity value has increased even for a small defect size of 250 mm.

The overall velocity significantly increases to 66% in case of a maximum defect size of 1500 mm with respect to healthy bearing at 15 kg load. The spectrum of the vibration velocity signal in the low-frequency range was obtained to observe changes at the characteristic defect frequency of the bearing outer race due to defects in it. The characteristic defect frequency of the rolling element bearing outer race can be calculated by using following equation [3, 13]:fo(HZ)=N/2fr[1-bd/pdCosß) ……..(1)

Figure-2: Vibration Velocity of Healthy Bearings at 15 kg Load

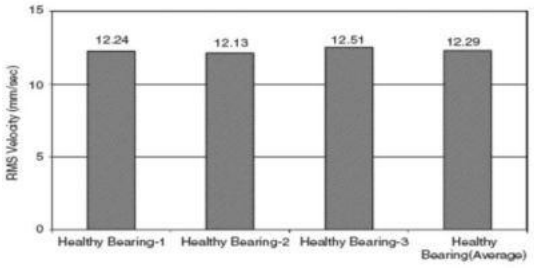


Fig. 2. Vibration velocity of healthy bearings at 15kg load.

Sources: Authors Compilation.

Figure-3: Overall Vibration Velocity with Outer Race Defects
Sources: Authors Compilation.

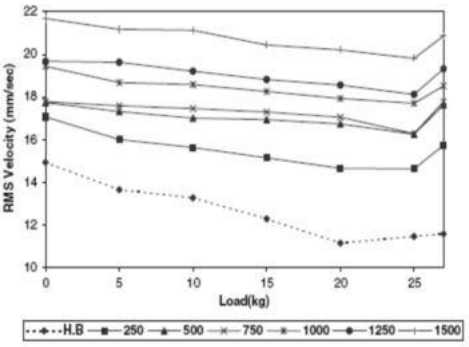


Fig. 3. Overall vibration velocity with outer race defects.

Where f_r is the shaft rotational frequency, N is number of balls, bd and pd are the ball diameter and pitch circle diameter, respectively, and b is the contact angle of the ball (with the races). For the shaft rotational frequency f_r= 24 Hz and a test bearing having nine balls of diameter 8.5mm and pitch circle diameter of 38.5mm with contact angle ß=0, the characteristic inner race defect frequency fo is found to be84.15 Hz. Velocity spectrum of one of the healthy bearings in the low-frequency range of 500 Hz at 15 kg load is shown in Fig. 4.From the spectrum of velocity, it has been observed that the peak occurs at fundamental frequency of shaft (i.e. at 24 Hz) and at twice the supply frequency (i.e. at 100 Hz) in the spectral component.

Figure-4: Spectrum of Vibration of Healthy Bearings

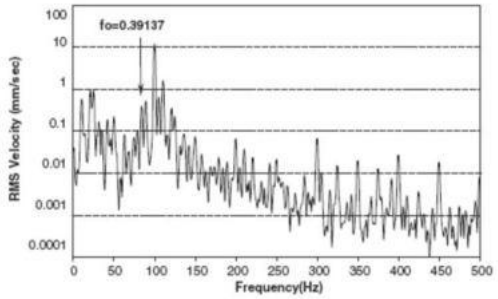


Fig. 4. Spectrum of vibration velocity of healthy bearing.

Sources: Authors Compilation.

Figure-5: Spectrum of Vibration for 1500µm Defect

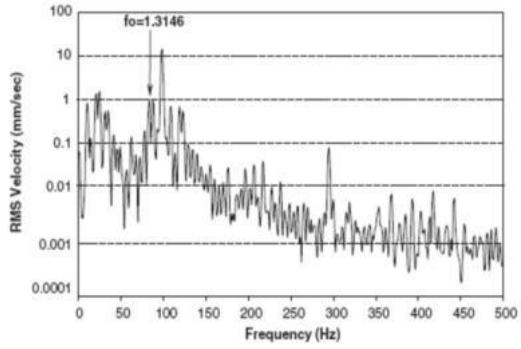


Fig. 5. Spectrum of vibration velocity for 1500µm defect.

Sources: Authors Compilation.

3.2. Stator Current Signals

Figure-6 shows the overall value of stator current amplitude comparison at 15 kg load for 0–20 kHz range of three healthy bearings and also the average overall stator current of three healthy bearings. From this chart also it is observed that the overall amplitudes are very much close tone another for the three healthy bearings. The overall stator current values were taken for all the defect sizes in the inner race of the bearing.

Figure-7 shows that the overall stator current values follow the same trend as that of the good bearing within crease in load. The overall stator current value has increased slightly even for a small defect size of 250 mm. From 250 to 1250 mm, the amplitude of overall stator current values increase continuously and the increase is much more for 1500 mm defect size. As mentioned in the literature, a defect in rolling element-bearing causes an increase in the overall RMS value of stator current for a known frequency range. Hence, the results obtained from the stator current correlates with the results reported in [4,15].

Figure-6: Overall Stator Current of Health Bearings at 15kg Load

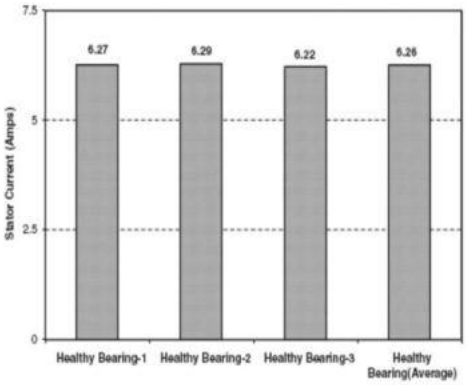


Fig. 6. Overall stator current of healthy bearings at 15kg load.

Sources: Authors Compilation.

Figure-7: Overall Stator Current with Outer Race Defects

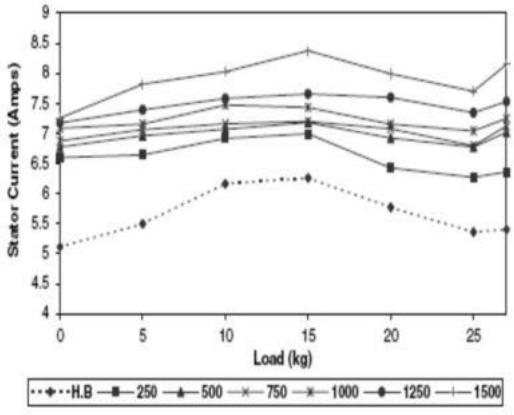


Fig. 7. Overall stator current with outer race defects.

Sources: Authors Compilation.

Thus, overall stator current has appreciably increased by 39.79% in case of maximum defect size of 1500 mm with respect to healthy bearing at15 kg load. The relationship of the bearing vibration to the stator current spectra can be determined by remembering that any air gap eccentricity produces anomalies in the air gap flux density [13].

Since ball bearings support the rotor, any bearing defect will produce a radial motion between the rotor and stator of the machine. The mechanical displacement resulting from damaged bearing causes the machine air gap to vary in a manner that can be described by combinations of rotating eccentricities moving in both directions [16].

Thus, bearing fault simulated in the inner race of the bearing may also cause rotor eccentricity, which is one of the common mechanical faults in the bearing. The rotor eccentricity in induction motor takes two forms, i.e. static eccentricity (where the rotor is displaced from the stator bore centre but still rotating upon its own axis) and dynamic eccentricity (where the rotor is still turning upon the stator bore centre but not on its own centre).

Eccentricity causes a force on the rotor that tries to pull the rotor even further from the stator bore centre. In the case of static eccentricity this is a steady pull in one direction. This makes the unbalanced magnetic pull (UMP) difficult to detect unless specialist experimental equipment is utilized, which is not possible for motors in service.

Dynamic eccentricity produces a UMP, which acts on the rotor and rotates at rotor rotational velocity. Both types of eccentricities cause excessive stressing of the machine and greatly increase bearing wear due to uneven magnetic pull produced that leads to variation of the sideband current magnitudes or predicted current harmonics in relation to vibration velocity.

Hence, any fault condition in the induction motor causes the magnetic field in the air gap of the machine to be non-uniform. It has been shown by Schoen [1] that these vibration frequencies reflect themselves in the current spectrum as f bng=[fe± mfv] …….. (2)

Where fe is the electrical supply frequency, m =1,2,3… is one of the characteristic vibration frequencies. A current signal of a single phase of stator current of induction motor and a vibration signal from a vibration sensor located at the bearing housing of induction motor for three good bearings were obtained. The corresponding current spectrum components in relation to vibrations for the supply frequency fe of 50 Hz and at characteristic outer race defect vibration frequency of 84.15 Hz are 34.15 and 134.15 Hz.

Fig. 8 shows the acquired current spectrum for the same low-frequency range of 500 Hz at 15 kg load as that of vibration velocity spectrum to verify the relationship between stator current and vibration velocity for healthy bearing. The spectrum of stator current in Fig. 8 indicates peak at supply frequency of 50 Hz in the current spectrum, whereas at twice the supply frequency in the velocity spectrum (Fig. 4) indicating the UMP even under normal operating condition. The corresponding current spectrum components in relation to vibrations were not significant at:

[f e _ f o] (i.e. at 34.15 Hz) and [fe+fo] (i.e. at 134.15 Hz) as shown in Fig. 8.

Figure-8: Spectrum of Stator Current of Healthy Bearings

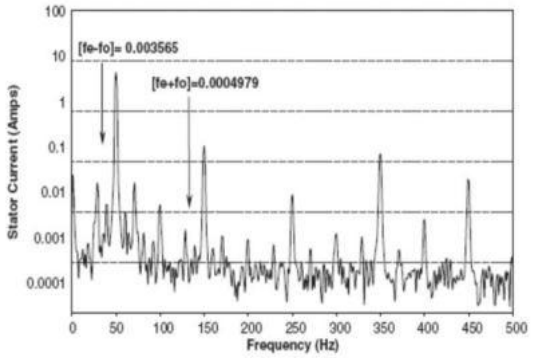


Fig. 8. Spectrum of stator current of healthy bearing.

Sources: Authors Compilation.

The rest of the peak component other than at 50 Hz present in the current spectrum occur at multiples of the supply frequency and these are caused due to saturation, winding distribution and supply voltage. The stator current spectrum of motor with the outer race defect of the bearing from 250 to 1500 mm in steps of 250 mm were obtained in the low frequency of 500 Hz for 15 kg load and the plot of 1500 mm is shown in Fig. 9.

Predicted current harmonics for outer race of the bearing relating vibration characteristic defect frequencies with the supply current frequency are compared with those of healthy bearing. For minimum defect size in the outer race of the bearing of motor, there was marginal increase in the amplitude of the predicted current harmonics component at [f e - f o] =34:15 Hz and [f e + f o] =134:15 Hz. However, significant increase in the amplitude of predicted current harmonics or vibration sideband current magnitudes is observed as the defect size increases as shown in Fig. 9. These results are comparable with results reported in [5, 13, and 17] for the outer race defect in the bearing.

3.3. Acoustic Emission Monitoring

In the present work, peak amplitudes of the AE signal of three healthy bearings from no load to full load were obtained. Peak amplitudes of the signal at 375 kHz were obtained and are expressed in dB with 0 dB corresponding to 1 V rams in the oscilloscope. The values of AE maximum peak amplitude of three healthy bearings are shown in Fig. 10.

The average value is 31.8 dB. Fig. 11 shows the AE maximum peak amplitude obtained from no load to full load for all defect sizes in the outer race of the bearing. It is observed that as the defect size increases the peak amplitude also increased. It has been observed from Fig. 11; the peak amplitude increased till 10 kg load and then decreases slightly with increase in load.

The range of AE maximum peak amplitudes from healthy bearing to maximum defect size are 31.4–67.8 dB at 10 kg load and 31.8–63.2 dB at15 kg load. In general the difference in maximum peak amplitude of healthy and smallest defect size is quite significant and makes it possible to detect the presence of a defect for diagnosis easier at the early stage in comparison with other condition monitoring techniques. There is an appreciable increase of 98.74% in case of maximum defect size with respect to healthy bearing at 15 kg load. Whereas, for 10 kg load, increase is 115.92% in case of maximum defect size with respect to healthy bearing.

3.4. Shock Pulse Method (SPM)

To neutralize the effect of rolling velocity on the measured value, the instrument was fed the shaft diameter of 25mm and rotational speed of 1440 rpm. After this SPM T2000 series calculates the initial shock value dBi, and displays maximum normalized shock pulse values.

Fig. 12 shows the comparison chart for maximum normalized value and average value of three healthy bearings. The values vary from 14 to 17 dB and are less than 20, indicating good bearing condition. Maximum normalized values (dBM) obtained for all defect sizes of250–1500 mm in the outer race of the bearing from no load to full load at constant motor speed are given in Fig. 13 in comparison with the average value of three healthy bearings.

It is seen that the levels for different bearings are more than that of healthy bearing and for 250–750 mm defect size; values of dBM are greater than20 and less than 35. For the defect size greater than 750 mm values of dBM are greater than 35. The dBM obtained indicates caution zone for the defect size of 250–750 mm and for the defect size above 750 mm dBM values obtained indicate the damaged bearing condition. Maximum normalized value as high as 50 were measured for the maximum defect size.

Figure-9: Spectrum of Stator Current for 1500µm Defect

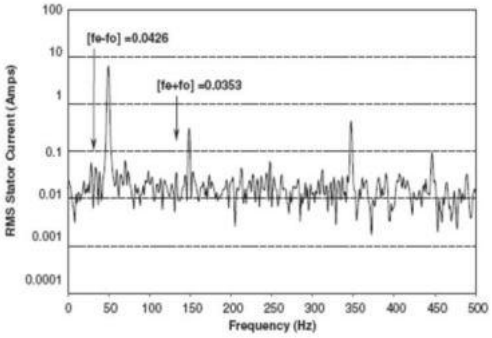


Fig. 9. Spectrum of stator current for 1500µm defect.

Sources: Authors Compilation.

Figure-10: Acoustic emission peak amplitude of Healthy Bearings at 15kg Load.

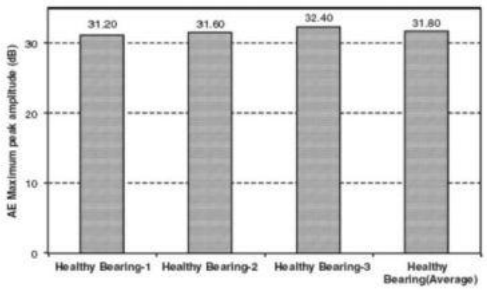


Fig. 10. Acoustic emission peak amplitude of healthy bearings at 15 kg load.

Sources: Authors Compilation.

Figure-11: Acoustic emission peak amplitude with outer race defects.

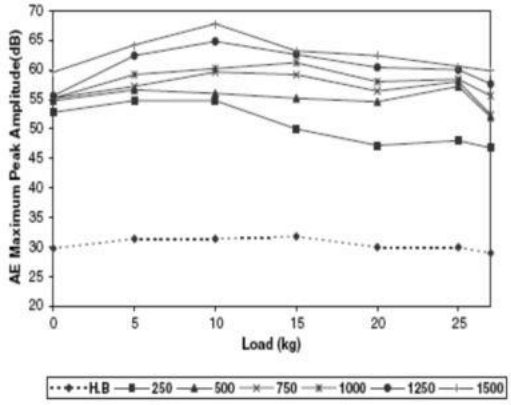


Fig. 11. Acoustic emission peak amplitude with outer race defects.

Sources: Authors Compilation.

IV. COMPARISON OF TECHNIQUES

Comparative study of different condition monitoring techniques has been done for minimum and maximum defect size in outer race of bearing. Fig. 12 shows the effectiveness of each technique in terms of percentage increase with respect to average value of healthy bearing, for the smallest defect size. As seen in Fig. 12, AE technique is the most effective technique followed by SPM.

The maximum normalized value of SPM is almost three times less effective as compared to AE technique. Overall vibration velocity and stator current come in the third and fourth place, respectively. Fig. 13 shows the same order in the effectiveness of each technique for maximum defect size.

However, stator current monitoring has the advantage that it requires minimum instruments and is sometimes referred to as sensor less technique.

Figure-12: Comparisoin of Condition Monitoring Techniques for Minimum Defect Size.

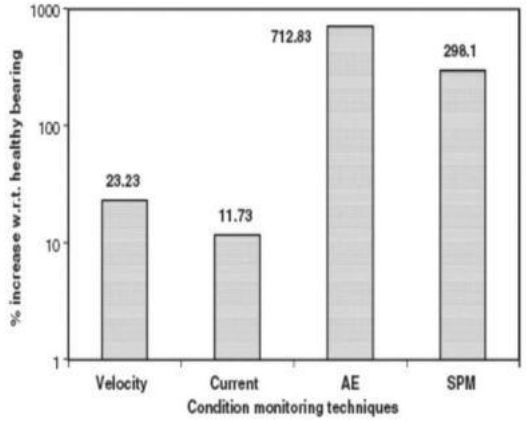


Fig. 14. Comparison of condition monitoring techniques for minimum defect size.

Sources: Authors Compilation.

Figure-13: Comparisoin of Condition Monitoring Techniques for Maximum Defect Size.

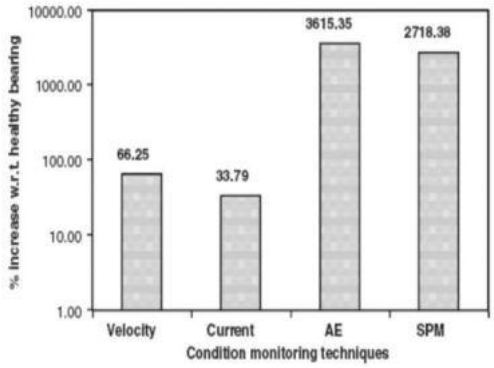


Fig. 15. Comparison of condition monitoring techniques for maximum defect size.

Sources: Authors Compilation.

IV. CONCLUSIONS

The vibration and stator current signal measurements performed on bearing of an induction motor are successful in detecting simulated defects in the outer race of the bearing. Current harmonics for bearing outer race defect characteristic vibration frequency has shown significant increase in the current spectrum components for maximum size of defect.

The AE and SPM measurement performed are very good in detecting the bearing defect. On comparing the results of good and defective bearing, it is observed that AE peak amplitude and shock pulse maximum normalized value level increase much more than other techniques as defect size increases. AE monitoring has proved to be the best technique. Stator current monitoring is perhaps the most cost-effective technique.

V. REFERENCES

[1] R., R. Schoen; B., K. Lin; T., G. Habetler; J., H. Schlag, and S., Farag, (2007). "An Unsupervised, On-Line System for Induction Motor Fault Detection Using Stator Current Monitoring", IEEE Transactions on Industry Applications, 31 (6), pp. 1280–1286.

[2] F., F. Albrechi; J., C. Apparatus; R., M. McCoy; E., L. Owen, and D., K. Sharma, (2009). "Assessment in the Reliability of Motors in Utility Application – Updated", IEEE Transactions on Energy Conversion, 1 (1), pp. 39–46.

[3] N., Tandon, and A., Choudhury, (1999). "A Review of the Vibration and Acoustic Measurement Methods for Detection of Defects in Rolling Element Bearings", Tribology International 32 (8), pp. 469–480.

[4] C., M. Riley; B., K. Lin; T., G. Hebelter, and R., R. Schoen, (2008). "A Method for Sensor-Less on Line Vibration Monitoring of Induction Machines", IEEE Transactions on Industry Applications, 34 (6), pp. 1240–1245.

[5] Y., Da-Mang, and P., James, (2000). "Intelligent Detection of Induction Motor Bearing Faults Using Current and Vibration Monitoring", Proceedings ''COMADEM 2000'', 3rd-8th December, pp. 461–470.

[6] T., Yoshioka, and T., Fuhjiwara, (2010). "Application of Acoustic Emission Technique to Detection of Rolling Element Bearing Failure", D.A. Dornfield (Ed.), Acoustic Emission Monitoring and Analysis in Manufacturing, ASME, New York, pp. 55–75.

[7] T. Yoshioka, Detection of rolling contact subsurface fatigue cracks using acoustic emission technique, Lubrication Engineering 49 (4)(2012) 303–308.

[8] T., Yoshioka, and M., Takeda, (2004). "Classification of Rolling Contact Fatigue Initiation Using Acoustic Emission Technique", LubricationEngineering, 51 (1), pp. 41–44.

[9] N., Tandon, and B., C. Nakra, (1990). "Defect Detection in Rolling Element Bearings by Acoustic Emission Method", Journal of Acoustic Emission, 9 (1), pp. 25–28.

[10] C., C. Tan, (1990). "Application of Acoustic Emission to the Detection of Bearing Failures", Proceedings "Tribology Conference", Brisbane, pp. 110–114.

[11] D., E. Butler, (2010). "The Shock Pulse Method for the Detection of Damaged Rolling Bearings", NDT International, 6 (2), pp. 92–95.

[12] I., E. Morando, (2008). "Measuring Shock Pulses is Ideal for Bearing Condition Monitoring", Pulp and Paper, 62 (12), pp. 96–98.

[13] R., R. Schoen; T., G. Habetler; F., Kamran, and R., G. Barthheld, (2005). "Motor Bearing Damage Detection Using Stator Current Monitoring, IEEE Transactions on Industry Applications, 31(6), pp. 1274–1279.

[14] N., Tandon, and B., C. Nakra, (2003). "Detection of Defects in Rolling Element Bearings by Vibration Monitoring", Journal of Institution of Engineers (India), Mechanical Engineering Division, 73, pp. 271–282.

[15] M., El Hachemi Benbouzid, (2000). "A Review of Induction Motor Signature Analysis as a Medium for Fault Detection", IEEE Transactions on Industrial Electronics, 47 (5), pp. 984–993.

[16] C., M. Riley; B., K. Lin; T., G. Habetler, and G., B. Kliman, (1999). "Stator Current Harmonics and Their Casual Vibrations: A Preliminary Investigationof Sensor Less Vibration Monitoring Applications", IEEE Transactions on Industry Applications, 35 (1).

[17] M., El Hachemi Benbouzid; M., Vieira, and C. Theyes, (2009). "Induction Motors' Faults Detection and Localization Using Stator Current Advancedsignal Processing Techniques", IEEE Transactions on Ower Electronics, 14 (1), pp. 14–22.

*****

MULTICOST PARAMETERS BASED AODV ROUTING PROTOCOL FOR EFFECTIVE TRANSMISSION IN WIRELESS AD HOC NETWORKS

D. Loganathan[62] P. Ramamoorthy[63]

ABSTRACT

A Wireless ad hoc network is a collection of wireless mobile nodes that dynamically form a network connection temporarily without any support of static infrastructure. Affected by mobility of nodes, routing protocols aretaking a vital role in wireless transmission. There are mainly three types of routing protocols in ad hoc networks which are proactive, reactive and hybrid. The protocol presents the mechanism which reduces route loops and confirms trustworthy message exchange. The Ad hoc On Demand Distance Vector (AODV) routing algorithm is a reactive routing protocol designed for ad hoc mobile nodes.Since the network is compatible with the conventional, modify the routing protocol with allowed parameters.Hop count, total interference, node link delay, residual energy of a node and the node transmission power are the cost parameters assigned for link and path of the ad hoc networks. These parameters are combined in different optimization function with respect to AODV for selecting the optimal path. This paper accesses the regular AODV and Multicost Parameters Based AODV (MPB-AODV) Routing Protocol to acquire dynamic network performance metrics like Packet Delivery Ratio and Throughput of wireless ad hoc networks.

KEYWORDS

Ad hoc Networks, AODV Routing Protocol, Multicost Parameters, Network Performance etc.

INTRODUCTION

Wireless ad hoc networks are paradigms for mobile communication in which mobile nodes are with dynamism and randomly located in such a manner that communication between nodes does not depend on any underlying static network infrastructure.

The communication medium is broadcast and the nodes in a mobile ad hoc network are typically portable mobile devices with inhibited resources, Such as power, computation aptitude and storage capacity. Since no fixed infrastructure or centralized administration is Available, these networks are self-organized and end-to-end communication may require routing information [3] via several intermediate nodes. The routing protocols are vital role and it has to adapt quickly to the repeated changes in the ad-hoc network topology. Ad-hoc routing protocols are categorized into following three types:

a) Proactive or Table driven routing protocols: This kind of routing protocols are retains the network topology information in routing tables contains a updated list of destinations and their routes by time to time swapping their routing information with nearby nodes. Routing information is usually flooded in the entire network. At any time a node wants a route to the destination it runs a suitable path finding algorithm on the topology information it retains. E.g. DSDV, CGSR, WRP.

b) Reactive or on demand routing protocols: These kinds of protocols are not maintaining topology information of the network, with the help of connection establishment process nodes can obtain necessary route when it is required. And therefore this type of protocols is not exchanging the routing information time to time. e.g. DSR, AODV, TORA.

c) Hybrid routing protocols: In this protocols both proactive and reactive routing advantages are combined. The routing is in the beginning established with certain proactively prospected routes then it serves the demand from additionally activated nodes through reactive flooding. E.g. HRPLS, ZRP, HWMP.

From the above DSDV, AODV, DSR and ZRP which have been proposed for providing communication among all the nodes in the network. Due to lack of infrastructure and the limited transmission range of a node in a mobile ad hoc network [11] a node has to rely on neighbor nodes to route a packet to the destination node. In specific, all network functions are based on the node cooperation.

Currently, routing protocols for mobile ad hoc network, such as the Dynamic Source Routing (DSR) and the Ad hoc On Demand Distance Vector Routing Protocol (AODV) [1], [15] are based on the assumption that all nodes will cooperate. And without node cooperation, in a wireless ad hoc network, no route can be established; no packet can be forwarded, let alone any network applications. However, cooperative behavior, such as forwarding other node's messages, cannot be taken for decided.

This paper focuses and analyzes through network simulation (NS2) which compares the quality of service metrics like Throughput, Delay and Packet Delivery ratio of both regular AODV and Multicost Parameters Based AODV (MPB-AODV).Hop count (h), total interference (I), node link delay (d), residual energy of a node (R) and the node transmission power (T) are the cost parameters [5],[6],[7],[16] assigned for link and path of the ad hoc networks.

These parameters are combined in different optimization function [12] with respect to various routing algorithm for selecting the optimal path. The similar research have already done in modified AODV [8] and implemented earlier. The simulation result shows that the Multicost Parameters Based AODV performs well in significant metrics of wireless network performance.

RELATED WORKS

One of the common implementation of source routing is Ad hoc On Demand Distance Vector. AODV is completely on demand and need not require constant information updates so as to construct and keep up the routes. There are two major services like route discovery and route maintenance is only appealed when a mobile node requests them.

[62] Research Scholar, Anna University Tamil Nadu, India, dlogu_morappur@yahoo.co.in
[63] Dean, (ECE), SNS College of Technology, Tamil Nadu, India, pdrudhra.2007@gmail.com

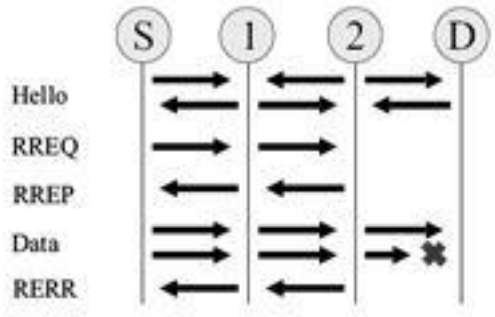
The AODV protocol have a both the features of reactive routing like DSR and proactive routing like DSDV techniques are used. It is similar to DSDV with usage of hop by hop routing, sequence numbers for route refresh ness and periodic beacons also similar to DSR in route discovery and maintenance. AODV can compact with any type of mobility rates and in different form of data traffic.

A numerous of research and many authors wanted to analyze the performance and behavior with regard to performance metrics of AODV and stated that to modify the AODV protocol [8] from that to improve the performance metrics like Throughput, Delay, Route latency and Packet Delivery ratio. And found that AODV protocol reveals much better QoS than DSDV and DSR.

AD HOC ON DEMAND DISTANCE VECTOR

AODV implements an exclusive technique to keep routing information in wireless ad hoc networks. AODV is on-demand routing protocol and basic essential for connectivity is to discover the routes to a mobile node through flooding of request messages.

Figure-1: AODV Routing Message



Sources: Authors Compilation.

Usually reactive protocols are never maintaining the routing information at the mobile nodes if no connectivity in the network. AODV uses old-style routing tables like one entry per destination node. Whereas DSR is maintains several routes cache entries for every destination node.

AODV finds route when node needs to communicate from source to destination and moreover assurance the loop-free routing. Every node is transmitting with other node through various wireless links and the nodes function as a router to route the data packets from one node to another node.

The AODV protocol mechanism as to send a message (Fig.1.), the data source starts a path-discovery process so as to discover the route. AODV routines sequence numbers retained at each destination to discover freshness of routing data and to avoid routing loops. The route request packet (RREQ) is flooded to the network and the transitional nodes record the neighbor from which they get the route request packet (RREQ) first, so as to establish inverse paths back to the source. When the RREQ reaches at the destination, it then directs back to a route reply (RREP) to the source node in reverse paths. AODV wants symmetric links; else the RREP may possibly not to reach the source and AODV might fail. And also, all the routing packets are bringing these sequence numbers.

A main feature of AODV is maintaining each node with timer-based states for deployment of individual routing table entries. If fail to use recent entry, the recent entry get expired in the routing table. A pair of predecessor nodes is maintained individually for the routing table entry, statingthat the pair of neighboring nodes to transmit the data packets.

In distinction with DSR, The Route Error Message (RERR) data packets in AODV [9] are projected to inform all sources using a link when a failure happens. Single-source shortest path Dijkstra's algorithm, computes of the shortest path from the source to every left behind vertices in the graph and find shortest path through Dijkstra's algorithm in AODV routing protocol.

MULTICOST PARAMETERS BASED AODV ROUTING PROTOCOL

A. AODV with Multicost Parameters

This section presents the improvement of the AODV protocol in order to strengthen the Packet Delivery Ratio, Average End-To-End Delay and Throughput in wireless ad hoc networks. Since the network is compatible with the conventional, modify the routing protocol with allowed parameters.The AODV routing protocol with other parameters are used to establish a route between the mobile nodes in the network as usual with performance metrics.

The experiment simulation shows that the network performance with respect to packet delivery ratio and throughput of the wireless ad hoc networks. The Multicost Parameters Based AODV (MPB-AODV) routing protocol with multicost parameters [10],[13],[14] used to when the origin node wishes to route a packet or a session to a given destination, a scalar cost optimization function 'f' is functional to the cost vectors of the non-dominated paths [2] leading to that destination, and the path that gives the minimum cost is chosen.

The optimization function f used depends on the QoS requirements of the session and may be different for different sessions. Since the optimization function does not need to be applied to each potential path for a given source-destination pair, but only to the set of non-dominated paths; this was proven in [7].

To be more exact (Eq.1), it can denote by:

$V(P) = (V_{1l}, V_{2l}, \ldots, V_{kl})$ the link cost vector of Link l, by $V(P) = (V_1, V_2, \ldots, V_k)$ the cost vector of the path P that contains of links $l = 1, 2\ldots L$ and by f (V) the optimizationfunction that has to be minimized in order to select the optimal path.

The cost vector $V(P) = (V_1, V_2, \ldots, V_k)$ of a path P containing of links $l = 1, 2\ldots L$ is then obtained from the cost vectors of the links that comprise it by spread on component-wise a monotonic associative operator Θ to each cost vector parameter(1):

$$V_m = \Theta_{l=1}^{L} V_{ml} \quad (1)$$

where $V_m = m^{th}$Parameter of the cost vector

B. SUM/MIN Energy-Half-Interference-Half Hop Multicost Algorithm

The optimization function is used as maximum representative of cost metrics in Various Energy-Half-Interference-Half Hop multicost algorithm:

– SUM/MIN Energy-Interference algorithm: The standard optimized through(Eq.2):

–

$$\min_P \frac{T_1(P) \cdot I_1(P)}{R(P).d(P)}, \quad (2)$$

It is likely to select paths that root slight total interference, use slight total transmission power, and permit through by nodes that have enormous residual energies.

– SUM/MIN Energy-Half-Interference algorithm: This cost function optimized is resembling with one used in the SUM/MIN Energy-Interference algorithm, but has a minor dependency with the interference metric(Eq.3):

–

$$\min_P \frac{T_\infty(P) \cdot \sqrt{I_\infty(P)}}{R(P).d(P)} \quad (3)$$

– SUM/MIN Energy-Interference-Half Hop algorithm: The optimization cost function is equivalent to the SUM/MIN Energy-Interference function, multiplied by $\sqrt{h(P)}$so as to depress, to a assured extent, the usage of long paths(Eq.4):

–

$$\min_P \frac{\sqrt{h(P)}.T_1(P) \cdot I_1(P)}{R(P).d(P)} \quad (4)$$

– SUM/MIN Energy-Half-Interference-Half Hop algorithm: The optimization function used is equivalent to that in the SUM/MIN Energy-Half-Interference algorithm(Eq.5), multiplied by $\sqrt{h(P)}$:

$$\min_P \frac{\sqrt{h(P)}.T_\infty(P) \cdot \sqrt{I_\infty(P)}}{R(P).d(P)}. \quad (5)$$

– MAX Interference algorithm: The cost function optimized is the maximum of the interferences of the links on the path(Eq.6):

$$\min_P I_1(P). \quad (6)$$

– MAX/MIN Energy-Half-Interference algorithm: The optimization function is close to that in the SUM/MIN Energy-Half-Interference algorithm, excluding that the transmission power and the interference are used as maximum representative instead of additive cost metrics(Eq.7):

$$\min_P \frac{T_\infty(P) \cdot \sqrt{I_\infty(P)}}{R(P).d(P)} \quad (7)$$

– MAX/MIN Energy-Half-Interference-Half Hop algorithm: The optimization function is also close to that in the SUM/MIN Energy-Half-Interference-Half Hop algorithm, excluding that the transmission power and the interference are used as maximum representative instead of additive cost metrics(Eq.8):

$$\min_P \frac{\sqrt{h(P)}.T_\infty(P) \cdot \sqrt{I_\infty(P)}}{R(P).d(P)} \quad (8)$$

Where, $h(P)$ = hop count of the path. $T_\infty(P)$ = maximum transmission power of the nodes on the path. I_∞= maximum interference of the path. $R(P)$ = residual energy of the path. $d(P)$ = delay link of the path.

Generally, the number of different non-dominated paths depends on the number of parameters in the cost vector, and on the type of operators used for calculating a path's cost vector from the establishes links' cost vectors. The cost parameters, h, d, are additive metrics, while R, T_∞ and $I_\infty$are concave (restrictive or maximum representative).

Based on [13], if the cost vector comprises at most one additive metric (other than the hop count), then the algorithm is polynomial, individually of the number of the restrictive (that use the minimization operator) and maximum representative (that use the maximization operator) metrics.

If the cost vector comprises two or more additive metrics (other than the hop count) then the algorithm is exponential. The complication considerations make some (polynomial) algorithms interesting even though they underperform some other (exponential) algorithms. As a result the SUM/MIN (Energy-Interference and Mixed) algorithm (Eq.5) is exponential and perform better than MAX/MIN Energy-Interference and Mixed) algorithm.

C. Simulation Model

Actually, most of the presently available devices afford a limited set of possible power levels of transmission power control. Assume that the transmission power can take sequence values are mainly made for the simulations, and is not essential by the algorithms themselves. To calculate the minimum transmission power for communicate among two nodes with some distance d can use the following equation mentioned below in Eq. 9.

$$P_r(d) = \frac{P_t \cdot G_t \cdot G_r \cdot \lambda^2}{(4\pi)^2 \cdot d^a \cdot L} \quad (9)$$

Where, P_r is the received power signal, P_t is the transmitted power signal G_t is the gain the senders antenna, G_r is the gains of receivers antenna, $L \geq 1$ the system loss and λ the wavelength used. In the calculations, assume $L = 1$ and $G_t = 1$, $G_r = 1$.

The parameter 'a' is the path loss constant and is usually in between 2 and 4 subject to the wireless channel as well as assume a = 2, corresponding to the Free Space transmission model.

PERFORMANCE COMPARISON

A. Simulation

Simulations play a dynamic role in the development and testing of ad hoc networking protocols. However, the simulation of large networks is still a tedious task that consumes a lot of computing power, memory, and time.

Table-1: Simulation Parameters for Node Mobility

PARAMETER	VALUE
Routing Protocol	AODV, Modified AODV
MAC Layer	802.11
Terrain Size	600m x 600m
No.of Nodes	50
Pause Time	10 seconds
Mobility Model	Random Mobility Model
Packet Size	1500B
Bandwidth	11MB
Frequency	2.472GHz
Speed	1-20 m/s
Simulation Time	100 s
Application Layer	FTP

Sources: Authors Compilation.

The changes were made to the implementation of AODV written for NS2. A 50 nodes network in a field size of 600m x 600m was used. The mobility model used was random waypoint in a square/rectangular field.

In random waypoint, each node starts its journey from its current location to a random location within the field. The speed is randomly chosen to be between 1-20 M/sec. The pause time is set to 10 seconds and to set the simulation time is 600 seconds.

Once the destination is reached, another random destination is embattled after a specified pause. Used here 10-second pause time, which results in unbroken node mobility in our simulations. The simulation parameters used for the experimental set are shown in Table-1.

However, in practice, found that the running times of the non-polynomial algorithms were also acceptable, at least for the network sizes used in the simulations. In all cases, the algorithms first find cost parameters (h, T, I, d, R), and then use the corresponding optimization function f (h, T, I, d, R) to select the optimal path with respect to AODV.

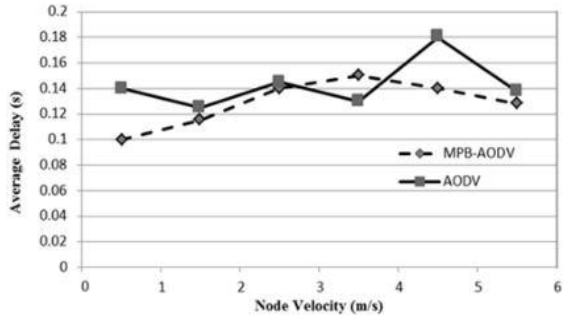
In other words the computation of algorithm and the AODV routing path is done at the end in a way proposed. The function to be optimized at the last step may depend on the QoS requirements of the user.

B. Proposed Improvements

The following performance metrics are conferred with AODV and Multicost Parameters Based AODV (MPB-AODV).Packet Delivery Ratio/ Packet Delivery: Throughput and Packet delivery ratio is calculated by dividing the number of packets received by the destination through the number of packets originated.

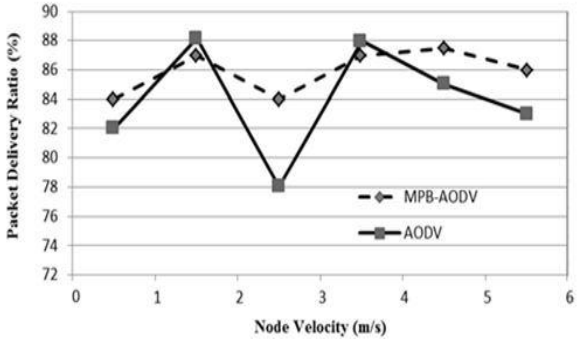
It specifies the packet loss rate, which limits the maximum throughput of the network. The better the delivery ratio, the more complete and correct is the routing protocol.

Figure-2: Average Delay When Node Velocity Increases



Sources: Authors Compilation.

Figure-3: Packet Delivery Ratio When Node Velocity Varies

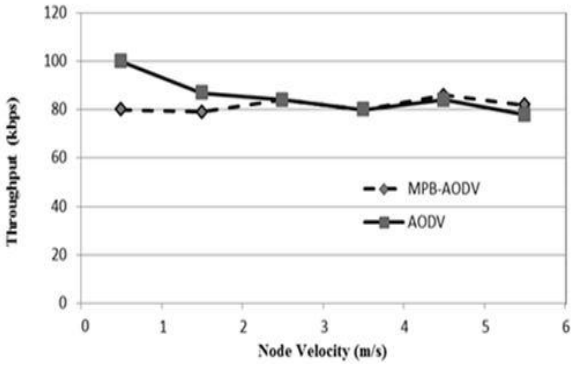


Sources: Authors Compilation.

Average End-To-End Delay: Average End-to-End delay (seconds) is the average time it takes a data packet to reach the destination. This metric is calculated by subtracting "time at which first packet was transmitted by source" from "time at which first data packet arrived to destination". This includes all possible delays affected by buffering during route discovery latency, queuing at the interface queue, retransmission delays at the MAC [4], propagation and transfer times.

Throughput: The throughput is defined as the total amount of data a receiver receives from the sender divided by the time it takes for the receiver to get the last packet. The throughput is measured in bits per second (bit/s or bps).

Figure-4: Throughput When Node Velocity Varies



Sources: Authors Compilation.

Throughput is the ratio of the total amount of data that reaches a receiver from a sender to the time it takes or the receiver to get the last packet. Packet Delivery Ratio/ Packet Delivery and Throughput are major importance of the performance of the network. The projected improvements in this paper to construct the enhancements in routing protocol of AODV with multicost parameters.

The Multicost Parameters Based AODV (MPB-AODV), Where the cost parameters of multicost algorithm h, d, R, T_∞ and I_∞ are carefully examined with AODV protocol and are combined in various optimization functions only at the end to improve the Packet Delivery Ratio and Throughput in wireless ad hoc networks. In AODV, when a host wants a route to another host, the route request packet (RREQ) is flooded to the network and the transitional nodes record the neighbor from which they get the route request packet (RREQ) first, so as to establish inverse paths back to the source. Simulation graph plotted on the different scales to best show the effects of varying packet delivery ratio, average delay and throughput in AODV and Multicost Parameters Based AODV. Based on the simulation results, Multicost Parameters Based AODV has the dynamic throughput Fig.4, a worthy packet delivery ratio Fig.3and uniformly good of average end to end delay Fig.2. Henceforward, Multicost Parameters Based AODV shows better performance with respect to throughput among basic AODV protocol in wireless ad hoc network in limited load density.

CONCLUSIONS

In this paper, the performance of the wireless routing protocols such as AODV and Multicost Parameters Based AODV (MPB-AODV) was analyzed using NS-2 Simulator. Deliberated complete simulation results of throughput, average delay and packet delivery ratio over the routing protocols AODV and Modified AODV by varying node velocity and simulation time. Data packet exchange will increase each time network topology changes since AODV protocol maintaining each node with timer-based states regarding deployment of individual routing table entries. Though Comparing Modified AODV protocol (MPB-AODV) with basic AODV, it performs better in case of packet delivery ratio but it performs slowly down in terms of throughput when increases node velocity in the network. Overall, Multicost Parameters Based AODV protocol outperforms is better because it has high packet delivery ratio and throughput when nodes have high mobility and considering with the Energy-Interference multicost algorithm.

REFERENCES

[1] Ad hoc On-Demand Distance Vector (AODV) Routing draft-ietf-manet-aodv-08.txt., March, 2001.

[2] Banerjee, S., and Misra, A., (2002). "Minimum Energy Paths For Reliable Communication In Multi-Hop Wireless Networks", ACM International Symposium on Mobile ad hoc Networking and Computing.

[3] C., Siva Ram Murthy, and B., S. Manoj, (2004). "Ad Hoc Wireless Networks: Architectures and Protocols", Prentice Hall.

[4] Samal, Chandra Kant, (2010. "TCP Performance through Simulation and Testbed in Multi-hop Mobile Ad hoc Network", IJCNC, Volume: 2, Number: 4, July.

[5] Ephremides, A., (2002). Energy Concerns in Wireless Networks", IEEE Wireless Communications, 9(4), pp. 48–59.

[6] Gupta, N., and Das, S. R., (2002). Energy-Aware On-Demand Routing For Mobile Ad Hoc Networks", International workshop on Distributed Computing, Mobile and Wireless Computing, pp. 164–173, New York: Springer.

[7] Gutierrez, F. J.; Varvarigos, E., and Vassiliadis, S., (2000) "Multicost Routing in Max-Min Fair Share Networks", Annual Allerton Conference on Communication, Control and Computing.

[8] Humaira, Nishat; Vamsi, Krishna K.; Dr. D., Srinivasa Rao, and Shakeel, Ahmed, (2011). "Performance Evaluation of On Demand Routing Protocols AODV and Modified AODV in MANETS", International Journal of Distributed and Parallel Systems", January, Volume: 2, Number: 1.

[9] Humaira, Nishat; Vamsi, Krishna K.; Dr. D., Srinivasa Rao, and Shakeel, Ahmed, (2011). "Performance Evaluation of On Demand Routing Protocols AODV and Modified AODV (R-AODV) in MANETS", International Journal of Distributed and Parallel Systems (IJDPS), January, Volume: 2, Number: 1.

[10] Karagiorgas, N.; Kokkinos, P.; Papageorgiou, C., and Varvarigos, E., (2007). "Routing in Wireless adhoc Networks with Variable Transmission Power", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC).

[11] Li, Y., and Ephremides, A., (2007). "Joint Scheduling, Power Control and Routing Algorithm for ad hoc Wireless Networks", Ad Hoc Networks, 5(7), pp. 959–973.

[12] Loganathan, D., and Ramamoorthy, P., (2012). "Multicost Routing Approach in Wireless Ad hoc Networks", Journal of Computer Science, 8 (10): 1739-1748, pp. 3773- 3778. Science Publications.

[13] Nikos, Karagiorgas; Panagiotis, Kokkinos; Christos Papageorgiou, and Emmanouel, Varvarigos, (2010). "Joint Multi-Cost Routing and Power Control in Wireless Ad Hoc Networks", Springer Science Business Media, LLC 2010.

[14] Loganathan, D., and Ramamoorthy, P., (2012). "Multicost Energy-Interference Algorithm for Effective Transmission in Wireless Ad hoc Networks", European Journal of Scientific Research, Volume: 84, Number: 4, pp. 572-586.

[15] Perkins, C., (1997). Ad hoc on demand distance vector (aodv) routing, Internet Draft, draft-ietf-manet-aodv-00.txt.

[16] Shah, R., and Rabaey, J., (2002). "Energy Aware Routing for Low Energy ad hoc Sensor Networks", IEEE Wireless Communications And Networking Conference.

*****

A NOVAL APPROACH TO CRIME DETECTION USING COMMUNAL AND SPIKE DETECTION

Jeya Christy C.[64] N. Mookhambika[65]

ABSTRACT

Identity crime has become prominent because there is so much real identity data available on the Web, and confidential data accessible through unsecured mailboxes. It has also become easy for perpetrators to hide their true identities. This can happen in a myriad of insurance, credit, and telecommunications fraud, as well as other more serious crimes. In addition to this, identity crime is prevalent and costly in developed countries that do not have nationally registered identity numbers. Credit card fraud is an element of identity fraud. It can have far reaching effects, since the information on the card can be used to perpetrate other types of identity theft crimes. From using the signature on the back of a card that is stolen, to loaning a credit card to a friend or family member can cause someone to obtain what they need to open other credit card accounts or bank accounts in the victim's name. Credit applications are Internet or paper-based forms with written requests by potential customers for credit cards, mortgage loans, and personal loans. Credit application fraud is a specific case of identity crime, involving synthetic identity fraud and real identity theft.

This paper proposes new multilayered detection system complemented with two additional layers: 1) Communal Detection (CD), and 2) Spike Detection (SD). CD finds real social relationships to reduce the suspicion score, and is tamper resistant to synthetic social relationships. It is the whitelist-oriented approach on a fixed set of attributes. SD finds spikes in duplicates to increase the suspicion score, and is probe-resistant for attributes. It is the attribute-oriented approach on a variable-size set of attributes.

KEYWORDS

Credit Card, Fraud, Identity, Communal Detection (CD), Spike Detection (SD) etc.

RELATED WORKS

The Existing System use business rules and scorecards. In Australia, one business rule is the hundred-point physical identity check test which requires the applicant to provide sufficient point-weighted identity documents face-to-face. They must add up to at least 100 points, where a passport is worth 70 points. Another business rule is to contact (or investigate) the applicant over the telephone or Internet.

The business rules and scorecards, and known fraud matching have limitations. Another existing is known as fraud matching. Here, known frauds are complete applications which were confirmed to have the intent to defraud and usually periodically recorded into a blacklist. Subsequently, the applications are matched against the blacklist due to long time delays, in days or months, for fraud to reveal itself, and be reported and recorded. This provides a window of opportunity for fraudsters. Second, recording of frauds is highly manual. This means known frauds can be incorrect, expensive, and difficult to obtain, and have the potential of breaching privacy.

PROPOSED SYSTEM

The Proposed System proposes a new multilayered detection system complemented with two additional layers: communal detection (CD) and spike detection (SD). CD finds real social relationships to reduce the suspicion score, and is tamper resistant to synthetic social relationships. It is the white list-oriented approach on a fixed set of attributes.

SD finds spikes in duplicates to increase the suspicion score, and is probe-resistant for attributes. It is the attribute-oriented approach on a variable-size set of attributes. Together, CD and SD can detect more types of attacks, better account for changing legal behavior, and remove the redundant attributes.

Communal Detection

The CD algorithm works in real time by giving scores when there are exact or similar matches between categorical data. This section motivates the need for CD and its adaptive approach.

Suppose there were two credit card applications that provided the same postal address, home phone number, and date of birth, but one stated the applicant's name to be John Smith, and the other stated the applicant's name to be Joan Smith. These applications could be interpreted in three ways:

1. Either it is a fraudster attempting to obtain multiple credit cards using near duplicated data.
2. Possibly there are twins living in the same house who both are applying for a credit card, or,
3. It can be the same person applying twice, and there is a typographical error of one character in the first name.

Spike Detection

CD has a fundamental weakness in its attribute threshold. Specifically, CD must match at least three values for our data set. With less than three matched values, our white list does not contain real social relationships because some values, such as given name and unit number, are not unique identifiers. The fraudster can duplicate one or two important values which CD cannot detect. Before proceeding with a description of SD, it is necessary to reinforce that CD finds real social relationships to reduce the suspicion score, and is tamper resistant to synthetic social relationships. SD finds spikes to increase the suspicion score, and is probe resistant for attributes. Probe resistance reduces the chances a fraudster will discover attributes used in the SD score calculation. It is the attribute-oriented approach on a variable-size set of attributes.

MODULES

The main modules are: a) Admin Module, b) User Module, c) Identify Crime User, and d) Online Shopping.

[64] Student, Final Year, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, christy355@gmail.com
[65] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

1. Admin Module

In this module, the admin can add product details (product name, price, validity etc.) based on the category likes mobiles, computers, laptops etc., and maintain the product details. The user enter their credit card details, the credit card is valid by Communal Detection and Spike Detection. If the card details is valid, the user can purchase their items else it report to the admin as "fraud transaction".

2. User Module

The user can select purchasing products displayed in the home page or search the product using keyword or based on category. Then user can purchase the product using credit/debit card. To purchase, the user need to provide the following details like (credit card number, card holder name, date of birth, credit card provider). If the credit card is valid the user is allowed to purchase the product else it reports to the admin as "fraud transaction occurs".

3. Identify Crime User

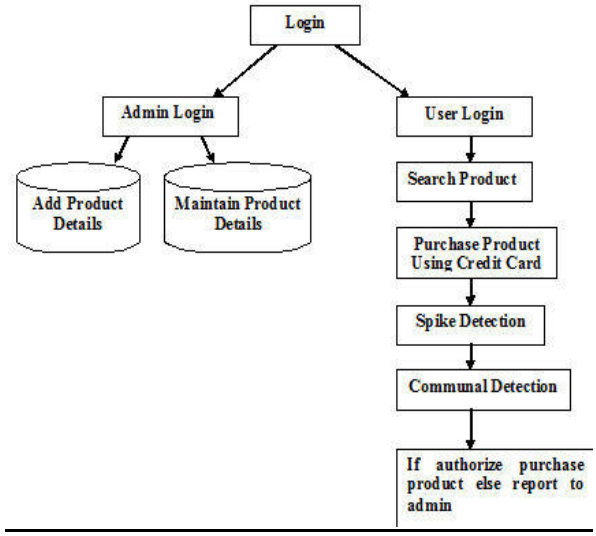
In this module we are using a real dataset: https://sites.google.com/site/cliftonphua/communal-fraud-scoring-data.zip. There are about 30 raw attributes such as personal names, addresses, telephone numbers, driver license numbers (or SSN), DoB, and other identity attributes (but no link attribute). Only 19 of the most important identity attributes are selected in table format as Training Dataset and Test Dataset. In the Dataset we identify the fraud transaction name by user name. The graph displays the user fraud percentage across months and measure weight for Test dataset user.

4. **Online Shopping**

In module, we developed a website for online shopping. The user can purchase a products using credit card. If the fraud user uses their credit card to purchase item, bank identify the fraud user using weight and graph of the user in the bank database.

DATA FLOW DIAGRAM

Figure-1



Sources: Authors Compilation.

Inputs
v_i (current application)
W number of v_j (moving window)
$\Re_{x,link-type}$ (link-types in current whitelist)
$T_{similarity}$ (string similarity threshold)
$T_{attribute}$ (attribute threshold)
η (exact duplicate filter)
α (exponential smoothing factor)
T_{input} (input size threshold)
SoA (State-of-Alert)

Outputs
$S(v_i)$ (suspicion score)
Same or new parameter value
New whitelist

CD algorithm

Step 1: Multi-attribute link [match v_i against W number of v_j to determine if a single attribute exceeds $T_{similarity}$; and create multi-attribute links if near duplicates' similarity exceeds $T_{attribute}$ or an exact duplicates' time difference exceeds η]

Step 2: Single-link score [calculate single-link score by matching Step 1's multi-attribute links against $\Re_{x,link-type}$]

Step 3: Single-link average previous score [calculate average previous scores from Step 1's linked previous applications]

Step 4: Multiple-links score [calculate $S(v_i)$ based on weighted average (using α) of Step 2's link scores and Step 3's average previous scores]

Step 5: Parameter's value change [determine same or new parameter value through SoA (for example, by comparing input size against T_{input}) at end of $u_{x,y}$]

Step 6: Whitelist change [determine new whitelist at end of g_x]

Inputs
v_i (current application)
W number of v_j (moving window)
t (current step)
$T_{similarity}$ (string similarity threshold)
θ (time difference filter)
α (exponential smoothing factor)

Outputs
$S(v_i)$ (suspicion score)
w_k (attribute weight)

SD algorithm

Step 1: Single-step scaled counts [match v_i against W number of v_j to determine if a single value exceeds $T_{similarity}$ and its time difference exceeds θ]

Step 2: Single-value spike detection [calculate current value's score based on weighted average (using α) of t Step 1's scaled matches]

Step 3: Multiple-values score [calculate $S(v_i)$ from Step 2's value scores and Step 4's w_k]

Step 4: SD attributes selection [determine w_k for SD at end of g_x]

Step 5: CD attribute weights change [determine w_k for CD at end of g_x]

CONCLUSIONS

We developed a real-time search (Online shopping) for patterns in multilayered to safeguard credit applications using communal detection (CD) and spike detection (SD).

CD finds real social relationships to reduce suspicion score, and is tamper resistant to synthetic social relationships. It is whitelist-oriented approach on a fixed set of attributes.

SD finds spikes in duplicates to increase the suspicion score, and is probe-resistant for attributes. It is attribute-oriented approach on a variable-size set of attributes. We are using a real dataset to identify credit card crime: https://sites.google.com/site/cliftonphua/communal-fraud-scoring-data.zip

FUTURE ENHANCEMENT

While in the experiments, CD and SD are updated after every period, it is not a true evaluation as the fraudsters do not get a chance to react and change their strategy in response to CD and SD as would occur if they were deployed in real life and consider the factor effectiveness, as scalability issues, extreme imbalanced class, and time constraints.

REFERENCES

[1] A., Bifet, and R., Kirkby, (2009). "Massive Online Analysis", Technical Manual, University of Waikato.

[2] R., Bolton, and D., Hand, (2001). "Unsupervised Profiling Methods for Fraud Detection", Statistical Science, Volume: 17, Number: 3, pp. 235-255.

[3] P., Brockett; R., Derrig; L., Golden; A., Levine, and M. Alpert, (2002). "Fraud Classification Using Principal Component Analysis of RIDITs", The Journal Risk and Insurance, Volume: 69, Number: 3, pp. 341-371, doi: 10.1111/1539-6975.00027.

[4] R., Caruana, and A., Niculescu-Mizil, (2004). "Data Mining in Metric Space: An Empirical Analysis of Supervised Learning Performance Criteria", Proceeding 10th ACM SIGKDD International Conference Knowledge Discovery and Data Mining (KDD '04), doi: 0.1145/1014052.1014063.

[5] P., Christen, and K., Goiser, (2007). "Quality and Complexity Measures for Data Linkage and Deduplication", Quality Measures in Data Mining, F. Guillet and H. Hamilton, eds., Volume: 43, Springer, doi: 10.1007/978-3-540-44918-8.

[6] C., Cortes; D., Pregibon, and C., Volinsky, (2003). "Computational Methods for Dynamic Graphs", Journal Computational and Graphical Statistics, Volume: 12, Number: 4, pp. 950-970, doi: 10.1198/1061860032742.

[7] Experian, (2008). "Experian Detect: Application Fraud Prevention System", Whitepaper, http://www.experian.com/products/pdf/experian_detect.pdf

*****

AN ONTOLOGY BASED ANATOMY APPROACH TO TEMPORAL TOPIC SUMMARIZATION

Thanuja P.[66] N. Mookhambika[67]

ABSTRACT

Searching the web is a vital role in human life in last two decades. The users either search for exact information or just surf topics which interest them. A topic is defined as a seminal event or activity along with all directly related events and activities. It is represented by a chronological sequence of documents published by different authors on the Internet.

An anatomy-based summarization method called Topic Summarization and Content Anatomy (TSCAN) was proposed to summarize the content of a temporal topic. TSCAN models the documents as a symmetric block association matrix, in which each block is a portion of a document, and treats each eigenvector of the matrix as a theme embedded in the topic. Temporal Similarity (TS) function is applied to generate the event dependencies and context similarity to form an evolution graph of the topic.

A unique feature of TSCAN is the introduction of the event segmentation process to extract the semantic construct event before summarization. In addition, we use an ontology based approach to improve the TSCAN. The idea of using ontology is to analyze documents and obtain semantic information. Before constructing symmetric block association matrix, we construct an ontology database for analyzing the main topics of the document.

After recognizing the main topics and determining their relative significance, we rank the paragraphs based on the relevance between main topics and each individual paragraph. Depending on the ranks, we decide preferred proportion of paragraphs as summary.

KEYWORDS

Topic Summarization and Content Anatomy (TSCAN), Temporal Similarity, Ontology, Anatomy etc.

INTRODUCTION

The phenomenal growth in the number of documents posted on the Internet provides an abundant source of information as an alternative to traditional media. While current technologies are efficient in searching for appropriate documents to satisfy keyword search requests, users still have difficulty assimilating needed knowledge from the overwhelming number of documents.

The situation is even more confusing if the desired knowledge is related to a temporal incident about which many independent authors have published documents based on various perspectives that, considered together, detail the development of the incident. To promote research on detecting and tracking incidents from Internet documents, the Defense Advanced Research Projects Agency (DARPA) initiated the Topic Detection and Tracking (TDT) project [1]. Its goal is to detect topics automatically and track related documents from several document streams, such as online news feeds. The TDT project has generated a great deal of interest due to importance and practical implications of the problem. For instance, the Google News service (see Fig. 1) employs TDT techniques to organize documents related to news topics from online news websites. While an effective TDT system can detect topics and track all related documents users cannot fully comprehend a topic unless they read many of the tracked documents.

Topic anatomy is an emerging text mining research paradigm that involves three major tasks: theme generation, event segmentation and summarization, and evolution graph construction. Generally, the content of a topic is comprised of several simultaneous themes, each representing an episode of the topic. The theme generation process tries to identify the themes of a topic from the related documents.

Figure-1: Google News service (http://news.google.com)

Sources: Authors Compilation.

In proposed system first, we collect vocabularies and synonyms. Then, we put those terms in the Data model of ontology. The first step of ontology based TSCAN approach is to compare the terms of documents with terms in the ontology. If the term does not exist in the ontology, we ignore it. Otherwise, we record the number of times the word appears in the ontology.

The ontology decomposes the specific domain into several objects for describing them. The determination of the way we describe objects and the formalism of representation depend on individual applications. The ontology is designed for analyzing and gathering the semantic information of a class of article. Assuming every document contains several subtopics; we use the ontology for identifying subtopics of document, and encode each of these possible subtopics by a non-overlapping portion

[66] Student, Final Year, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, jathanu47@gmail.com
[67] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

of the ontology. After selecting the blocks using ontology, we construct symmetric block association matrix.

CREATING ONTOLOGY

An ontology formed using protégé tool and by natural language processing. The terms in the document that needs to be summarized are matched against the entities present in the ontology to form a graph. This graph contains entities which represent the theme of the document. Each of the graph nodes is assigned a weight based on various factors. The largest connected component is then chosen to be the thematic graph that best represents the document content with respect to the ontology. Several graph centrality measures are used to further determine the importance of entities with respect to the document and scores are assigned to the nodes. These scores are then used to determine the most important sentences in the document which become part of the final summary.

TSCAN SYSTEM

In this section, we present our model and the methods used in the proposed topic anatomy system.

3.1 Topic Model

A topic is a real world incident that comprises one or more themes, which are related to a finer incident, a description, or a dialogue about a certain issue. During the lifespan of a topic, one theme may attract more attention than the others, and is thus reported by more documents. We define an event as a significant theme development that continues for a period of time. Naturally, all the events taken together form the storyline(s) of the topic. Although the events of a theme are temporally disjoint, they are considered semantically dependent in order to express the development of the theme. Moreover, events in different themes may be associated because of their temporal proximity and context similarity. The proposed method identifies themes and events from the topic's documents, and connects associated events to form the topic's evolution graph. In addition, the identified events are summarized to help readers better comprehend the storyline(s) of the topic.

A topic is represented explicitly by collection of chronologically ordered documents. In this study, we assume that documents are published in the same order as the events of the topic reported by independent authors, and that there is no inconsistency between the contents of documents. TSCAN decomposes each document into sequence of nonoverlapping blocks. A block can be several consecutive sentences, or one or more paragraphs. We define a block as w consecutive sentences.

For a topic, let $T=\{ t_1, t_2, \ldots, t_m\}$ be a set of stemmed vocabulary without stopwords [2]. The topic can then be described by an m×n term-block association matrix B in which the columns $\{b_1, b_2, \ldots, b_n\}$ represent the blocks decomposed chronologically from the topic documents. In other words, for any two blocks, b_i and b_j, if $i < j$, then either the document containing b_i was published before the document containing b_j, or b_i appears before b_j in the same document. The (i,j)-entry of B (denoted as $b_{i,j}$) is the weight of term i in block j, computed by using the well-known TF-IDF term weighting scheme [2].

3.2 Theme Generation

A matrix $A = B^TB$, called a block association matrix, is an n × n symmetric matrix in which the (i,j)-entry (denoted as $a_{i,j}$) is the inner product of columns i and j in matrix B. As a column of B is the term vector of a block, A represents the interblock association. Hence, entries with a large value imply a high correlation between the corresponding pair of blocks. A theme of a topic is regarded as an aggregated semantic profile of a collection of blocks, and can be represented as a vector $\underline{v}$ of dimension n, where each entry denotes the degree of correlation of a block to the theme. Given the constitution of a vector $\underline{v}$, $\underline{v}^T A\underline{v}$ computes the theme's association to the topic's content. The objective function in (1) of our theme generation process determines $\underline{v}$'s entry values so that the acquired theme is closely associated with the topic.

$$\max \underline{v}^T A\underline{v} \quad (1)$$
$$\text{s.t. } \underline{v}^T \underline{v} = 1. \quad (2)$$

Without specifying any constraint on $\underline{v}$, the objective function (1) becomes arbitrarily large with large entry values of $\underline{v}$. Constraint (2) limits the search space to the set of normalized vectors such that the following Lagrangian formula [3] can be used to solve (1) and (2).

$$Z(v,\lambda)= \underline{v}^T A\underline{v} + \lambda(1- \underline{v}^T \underline{v}). \quad (3)$$

To obtain the entry values of $\underline{v}$, let $\partial Z/\partial \underline{v} = \partial Z/\partial \lambda = 0$ as follows:

$$\partial Z/\partial \underline{v} = 2A\underline{v} - 2\lambda \underline{v} = 0 \quad (4)$$
$$\partial Z/\partial \lambda = 1 - \underline{v}^T \underline{v} = 0. \quad (5)$$

Equation (4) implies that $A\underline{v}=-\lambda\underline{v}$. In other words, $\underline{v}$ is a normalized eigenvector of A and λ is the corresponding eigenvalue. For any n × n square matrix, there are at most n eigenvectors [4]. In terms of nonlinear programming, (3) can have more than one stationary point [3].

3.3 Event Segmentation and Summarization

A theme v_j in V_L is a normalized eigenvector of dimension n, where the (i,j)-entry $v_{i,j}$ indicates the correlation between a block i and a theme j. As topic blocks are indexed chronologically, a sequence of entries in v_j with high values can be taken as a noteworthy event embedded in the theme, and valleys (i.e., a sequence of small values) in v_j may be event boundaries. However, according to definition of eigenvectors, the signs of entries in an eigenvector are invertible.

Moreover, Kleinberg [5] and Nicholas and Dahlberg [6] showed that both the positive and negative entries of an eigenvector contain meaningful semantics for describing a certain concept embedded in a document corpus; and the amplitude of an entry determines the degree of its correlation to the concept. Note that the tasks of our event segmentation and speech endpoint detection are similar in that they both try to identify important segments of sequential data. In addition, it is the amplitude of sequential data that determines the data's importance. Therefore, we adopt Rabiner and Sambur's R-S endpoint detection algorithm [7] for event segmentation.

To segment events, the R-S algorithm examines the amplitude variation of an eigenvector to find the endpoints that partition the theme into a set of significant events. In the R-S algorithm, every block in an eigenvector has an energy value. To calculate

the energy, we adopt the square sum scheme, which has proved effective in detecting endpoints in noisy speech environments [14]. Note that the summary block might not be the one with the largest energy value because of the averaging effect of the sliding window. Another interesting by-product of the above method is that the produced energy contour also describes the activeness trend and evolution of a theme. A unique feature of our summarization approach is the introduction of the event segmentation process to extract the semantic construct "event" before summarization. Most existing generic summarization approaches [8], [9], [10], [11], [12], [13] try to cover diverse themes in document summaries; however, our method further describes the development of themes via summarized events to help users comprehend a topic's storylines.

3.4 Evolution Graph Construction

An evolution graph connects themes and events to present the storylines of a topic. Let $X=\{e_1, e_2, \ldots, e_x\}$ be a set of events in a topic. For each event e_k, let $e_k.ev$ [1, L] denote the theme index of the event, and let $< e_k.fb, e_k.lb >$ be the event's timestamp, where $e_k.fb$ and $e_k.lb$ are the indexes of the first and last blocks, respectively; and $|e_k| = 1 + e_k.lb - e_k.fb$ is the temporal length of e_k. The topic evolution graph $G=(X,E)$ is a directed acyclic graph, where X represents the set of nodes and $E=\{(e_i,e_j)\}$ is the set of directed edges. An edge (e_i,e_j) specifies that event j is a consequent event of event i, which satisfies the constraint $e_j.fb > e_i.fb$.

Automatic induction of event dependencies is often difficult due to the lack of sufficient domain knowledge and effective knowledge induction mechanisms [15]. However, as event dependencies usually involve similar contextual information, such as the same locations and person names, they can be identified through word usage analysis. Our approach, which is based on this rationale, involves two procedures. First, we link events segmented from the same theme sequentially to reflect the theme's development. Then, we use a temporal similarity function to capture the dependencies of events in different themes. For two events, e_i and e_j, belonging to different themes, where $e_j.fb > e_i.fb$, we calculate their temporal similarity by:

$$TS(e_i,e_j)=TW(e_i,e_j)*cosine(e_i.cv,e_j.cv)$$

Where, the cosine function returns the cosine similarity between the centroid vectors of the events. The temporal weight (TW) function weights the similarity based on the temporal difference between the events. If the temporal similarity is above a predefined threshold, we deem e_j a consequence of e_i and construct a link between them.

PERFORMANCE EVALUATIONS

4.1 Data Corpus

In [16], two case studies using the official TDT topics are provided to demonstrate that the evolution graphs constructed by TSCAN can extract the themes, events, and event dependencies of the examined topics successfully. In this research, we evaluate our anatomy-based summarization technique by comparing the derived summaries with those of several text summarization methods. We use the official TDT4 topics for the performance evaluations. The Linguistic Data Consortium has compiled a series of TDT corpora for the annual TDT contests. The TDT4 corpus comprises 28,390 English news documents from eight well-known news agencies for the period 1st October 2000 to 31st January 2001. Among them, 70 news events with 1,926 related documents were labeled by NIST annotators for various TDT evaluation tasks. The annotators also composed factual descriptions of the topics, which are regarded as human-composed reference summaries for summarization evaluations. Although Document Understanding Conferences 1 (DUC) also uses TDT topics for summarization contests, the average size of the topics is only 10 documents, which is too small for the purpose of topic anatomy. We therefore select 26 TDT4 topics, each containing more than 20 documents, for evaluation.

4.2 Summarization Evaluations

We compare the summarization performance of TSCAN with the following six well-known summarization methods:

1. The forward method which generates summaries by extracting the initial blocks of a topic.
2. The backward method, which extracts summaries from the end blocks of a topic. This is frequently used as the baseline method in DUC contests.
3. The SVD method [8], which composes summaries by extracting the blocks with the largest entry value in singular vectors. Note that the result derived by the SVD method is identical to that of the graph based summarization method [12].
4. The K-means method [9], which compiles summaries by selecting the most salient blocks of the resulting K clusters. Generally, this method's performance depends on the quality of the initial clusters. In our experiments, to ensure fair comparison of the Kmeans method, we use the best result from 50 randomly selected initial clusters for evaluation.
5. The temporal summary (TS) method [11], where we adopt the useful$_2$ and novel$_1$ techniques proposed by the authors to compute the informativeness score of a topic block. We do not adopt the novel$_2$ technique because the authors have shown that the performance difference between using novel$_1$ and using novel$_2$ is not significant. In addition, novel$_2$ requires a training corpus to derive an appropriate number of clusters (i.e., parameter m), but the training corpus is not available.
6. The frequent content word (FCW) method [13], which constructs summaries by selecting blocks with frequent terms. This method's performance is comparable to that of state-of-the-art summarization methods. In addition, we adopt Nenkova et al.'s context adjustment technique to increase the summary diversity.

We use four metrics, the summary-to-document content similarity (SDCS), average pairwise summary block similarity (APSBS), ROUGE [17], and execution time, to evaluate the above summarization methods. The SDCS metric compares the content coverage of a generated summary to that of the documents selected for summarization; APSBS measures the degree of content coherence of a summary; ROUGE considers the consistency between the content of a generated summary and that of a set of expert-composed reference summaries; and execution time measures the efficiency and scalability of the summarization methods. Naturally, it is preferable that the summary should be content-coherent and as close as possible to the reference summaries. It should also have a high degree of content diversity to cover all the important information in the summarized documents.

4.2.1 Summary-to-Document Content Similarity

Summary-to-document content similarity is defined as the average cosine similarity between an evaluated summary and the topic documents. Both components are represented by TF-IDF term vectors. A high similarity score implies that summary is representative of the topic and can effectively replace the original topic documents for various information retrieval tasks.

4.2.2 Average Pairwise Summary Block Similarity

Although our summaries are not as diverse as those of K-means method, they are more coherent. A popular measurement frequently used to judge the content coherence of a set of documents is the average pairwise document similarity [18]. The metric models documents as term vectors and computes the similarity between documents in terms of the cosine similarity. As the cosine similarity calculates the degree of word overlap between documents, a high average pairwise similarity indicates that the documents have a significant amount of word overlap; hence, the documents are content-coherent.

We employ the measurement to calculate the degree of content coherence of a summary. For an evaluated summary, which is represented by a set of summary blocks, we model each summary block as a TF-IDF term vector (i.e., b_i) and compute the average cosine similarity of all pairs of block vectors. We then average the average pairwise summary block similarity (APSBS) of the compared methods under various parameter settings. TSCAN achieves superior APSBS scores. The reason is that our summaries focus on events in the first few significant themes; therefore, summary blocks have similar contexts. By contrast, the summaries compiled by other approaches try to cover diverse themes, so they are less coherent than our summaries.

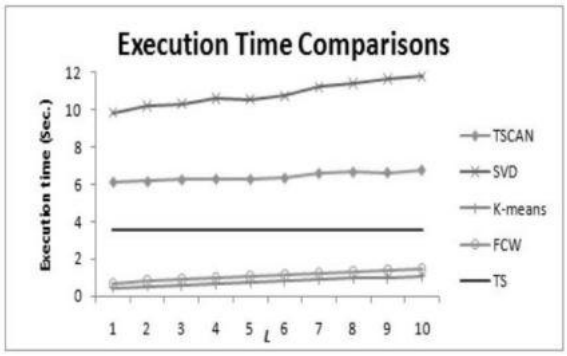
4.2.3 ROUGE Evaluations

ROUGE is a recall-oriented summary evaluation metric that is widely used in DUC contests. It measures the summarization performance by calculating the number of overlapping n-grams [18] between an evaluated summary and a set of reference summaries. Note that, in our study, the reference summaries are the topic explanations composed by NIST annotators. The ROUGE score is 1 if the evaluated summary is consistent with the reference summaries; and 0 otherwise. It has been shown that the results of comparisons based on ROUGE-1 and ROUGE-2 (i.e., unigram and bigram-overlap) are highly consistent with those derived by human evaluators. Therefore, we use ROUGE-1 and ROUGE-2 to evaluate the consistency of manual summaries derived by the compared methods.

4.2.4 Scalability and Time Comparisons

We evaluated execution time of the compared summarization methods on an AMD AthlonTM 64 Processor 3200++ PC with the Windows XP Service Pack 3 operating system and a 2 GB main memory. For each method, we recorded the time required to generate the summaries of the 26 evaluated topics under specific parameter setting. However, due to space limitations, we only show the average execution times of the methods under all parameter settings in Fig.2.

Figure-2: Comparison of Execution Times



Sources: Authors Compilation.

We do not show the execution times of the forward and backward methods because they do not need to weight each topic block to compile topic summaries. Therefore, their respective execution times are constant and irrelevant to the parameter settings. FCW is an iterative summarization method. In each iteration, it computes the weight of every block based on a topic model; therefore, its time complexity is O(n). K-means is also a linear algorithm [18].

Although the TS method computes the weight of topic blocks linearly, it needs to examine the content of previous blocks to compute a block's novelty; therefore, its time complexity is $O(n^2)$. The SVD method and our method spend much of the computation time calculating eigenvectors. Time complexity is $O(n^2I^2)$, where I is the number of eigenvectors to be computed .

The results in Figure-2 show that the linear summarization methods (i.e., Kmeans and FCW) run faster than the other methods. We believe that the longer execution time of the TS method is due to program implementation issues. Except for the TS method, the execution time of the compared methods generally increases as the size of the summary (i.e., L) increases. For the SVD method and TSCAN, a large L means that the methods need to examine a lot of eigenvectors to compile topic summaries; therefore, the execution time increases. For the K-means method and the FCW method, a large L increases the number of clusters (i.e., K) and the number of iterations required to extract summary blocks with frequent terms, respectively.

Hence, the methods' execution times also increase. It is noteworthy that the TS method's execution time is irrelevant to the size of the summary. This is because the method must weight all topic blocks irrespective of how many summary blocks are required to compile a topic summary.

4.2.5 Performance Review

To summarize, we evaluate the summarization performance of TSCAN in terms of content coverage, content coherence, consistency with expert-composed summaries, and execution time. The experiment results show that, as well as covering the core parts of evaluated topics, our summaries are content-coherent and consistent with expert composed reference summaries. The quality of our summaries is better than that of many well-known summarization methods, especially when the compression ratio is high. These results demonstrate that TSCAN can select representative sentences earlier than the compared methods when compiling topic summaries. In

resource-limited environments, such as when the network bandwidth is low, this property helps users capture key information about a topic. While our method's execution time is longer than that of many well-known summarization methods, the time required to compile the summaries of large topics is a few minutes at most; thus, it is feasible for practical text summarization systems. The improvements over the compared methods also emphasize the importance of temporal information (i.e., events) in temporal topic summarization. In addition to providing summary diversity, topic summarization methods should summarize the developments of significant themes that produce content-coherent summaries and are consistent with human-composed summaries.

CONCLUSIONS

Publishing activities on the Internet are now so prevalent that when a fresh news topic occurs, autonomous users may publish their opinions during the topic's life span. To help Internet users grasp the gist of a topic covered by a large number of topic documents, text summarization methods have been proposed to highlight the core information in the documents. Most summarization methods try to increase the diversity of summaries to cover all the important information in the original documents. However, when the documents to be summarized are related to an evolving topic, summarization methods should also consider the temporal properties of the topic in order to describe the development of storylines. By considering ontology attributes we are able to improve the semantic representation of a sentence's information content.

In this paper, we have presented a topic anatomy system called TSCAN, which extracts themes, events, and event summaries from topic documents. Moreover, the summarized events are associated by their semantic and temporal relationships, and presented graphically to form an evolution graph of the topic. We show that mapping provides a semantic representation of information content of sentences that improves summarization quality. Experiments based on official TDT topic demonstrate that TSCAN can produce highly representative summaries that correspond well to reference summaries composed by experts.

REFERENCES

[1] J., Allan; J., Carbonell; G., Doddington; J., Yamron, and Y., Yang, (1998). "Topic Detection and Tracking Pilot Study: Final Report", Proceedings US Defense Advanced Research Projects Agency (DARPA) Broadcast News Transcription and Understanding Workshop, pp. 194-218.

[2] R., Baeza-Yates, and B., Ribeiro-Neto, (1999). Modern Information Retrieval, Addison Wesley.

[3] W., L., (2004). Winston, Operations Research. Thomson.

[4] L., E. Spence; A., J. Insel, and S., H. Friedberg, (2000). "Elementary Linear Algebra, Matrix Approach", Prentice Hall.

[5] J., Kleinberg, (1998). "Authoritative Sources in a Hyperlinked Environment", Proceedings 9th Annual ACM-SIAM Symp. Discrete Algorithms, pp. 668-677.

[6] C., Nicholas, and R., Dahlberg, "(1998). "Spotting Topics with the Singular Value Decomposition", Proceeding 4th International Workshop Principles of Digital Document Processing, pp. 82-91.

[7] L., R. Rabiner, and M., R. Sambur, (1975). "An Algorithm for Determining the Endpoints for Isolated Utterances", the Bell System Technical Journal, February, Volume: 54, Number: 2, pp 297-315.

[8] Y., Gong and X., Liu, (2001). "Generic Text Summarization Using Relevance Measure and Latent Semantic Analysis", Proceeding 24th Annual International ACM SIGIR Conference Research and Development in Information Retrieval, pp. 19-25.

[9] T., Nomoto and Y., Matsumoto, (2001). "A New Approach to Unsupervised Text Summarization", Proceeding 24th Annual International ACM SIGIR Conference Research and Development in Information Retrieval, pp. 26-34.

[10] H., Zha, (2002). "Generic Summarization and Keyphrase Extraction Using Mutual Reinforcement Principle and Sentence Clustering", Proceeding 25th Annual International ACM SIGIR Conference Research and Development in Information Retrieval, pp. 113-120.

[11] J., Allan; R., Gupta, and V., Khandelwal, (2001). "Temporal Summaries of News Topic", Proceeding 24th Annual International ACM SIGIR Conference Research and Development in Information Retrieval, pp. 10-18.

[12] G., Erkan, and D., R. Radev, (2004). "LexRank: Graph-Based Centrality as Salience in Text Summarization", Journal Artificial Intelligence Research, Volume: 22, pp. 457-479.

[13] A., Nenkova; L., Vanderwende, and K., Mckeown, (2006). "A Compositional Context Sensitive Multi-Document Summarizer: Exploring Factors that Influence Summarization", Proceeding 29th Annual International ACM SIGIR Conference Research and Development in Information Retrieval, pp. 573-580.

[14] Y., Gong, and X., Liu, (2001). "Generic Text Summarization Using Relevance Measure and Latent Semantic Analysis", Proceeding 24th Annual International ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 19-25.

.

[15] R., Nallapati; A., Feng; F., Peng, and J., Allan, (2004). "Event Threading within News Topics", Proceeding 13th ACM International Confernce Information and Knowledge Management, pp. 446-453.

[16] C., C. Chen, and M., C. Chen, (2008). "TSCAN: A Novel Method for Topic Summarization and Content Anatomy", Proceeding 31st Annual International ACM SIGIR Conference Research and Development in Information Retrieval, pp. 579-586.

[17] C.Y. Lin and E. Hovy, (2003). "Automatic Evaluation of Summaries Using N-gram Co-Occurrence Statistics", Proceedings Conference North Am. Chapter of the Association for Computational Linguistics on Human Language Technology, Volume: 1, pp. 71-78.

[18] C., D. Manning; P., Raghavan, and H., Schutze, (2008). Introduction to Information Retrieval. Cambridge University Press.

*****

STUDY OF BARIUM TITANATE THIN FILMS BY SPIN COATING TECHNIQUE

D. Lakshmi[68] V. Meenapriya[69] J. Shanthi[70]

ABSTRACT

In recent years, barium titanate (BT) has attracted much attention to researchers, becoming one of the most investigated ferroelectric materials due to good electrical properties at room temperature, mechanical and chemical stability and the easiness in its preparation.

Deposition of pure barium titanate thin films by spin coating technique is reported in this study. The precursor solution is prepared by dissolving pure barium titanate powder in HNO3. The films were deposited on clean glass substrates by varying pH value, viscosity, time duration and rpm values.

The deposited films were analysed using powder X-ray diffraction technique, and it revealed sharp and intense peaks proving crystalline nature of material. Transmittance spectra of the films were also analysed by using UV-VIS spectrometry. Also, functional group identification was done using FTIR spectroscopy.

KEYWORDS

$BaTiO_3$, XRD, UV-VIS Spectrometry, FTIR etc.

INTRODUCTION

Ferroelectric materials such as barium titanate ($BaTiO_3$) possess a variety of useful properties that have potential to be exploited in thin-film devices. Device applications of ferroelectric films require that properties similar to those found in bulk be maintained in deposited.

This titanate is a ferroelectric ceramic material, with dominant photorefractive effect and piezoelectric properties. It is a good dielectric ceramic and can be used for capacitors. It functions as piezoelectric material for microphones and other transducers. Polycrystalline barium titanate displays positive temperature coefficient, making it a useful material for thermistors and self- regulating electric heating systems.

A variety of deposition techniques, such as rf sputtering, metal organic chemical vapor deposition MOCVD, molecular beam epitaxy MBE and pulsed laser deposition have been successfully used to synthesize $BaTiO_3$ (BTO) films. Among these, Sol–gel BTO coatings have also been previously prepared in order to study their chemical and micro-structural properties. [1,8] The sol–gel method presents several advantages for coatings processing; the control of composition, surface morphology engineering and low temperature processing, which allows the use of thermally fragile substrates, are its principal merits. [2,10] The purpose of this work was to prepare $BaTiO_3$ thin films using a spin coating method, which is a fast, simple, reproducible and low-temperature deposition technique. The effects of the spin speed annealing temperature on the physical properties of the films were investigated. The aim of the present was extended to analyse the structural, optical properties of $BaTiO_3$ thin films prepared by sol-gel spin coating process.

[68] Lectuer, VSB College of Engineering Technical Campus, Tamil Nadu, India, luxmi095@gmail.com
[69] Student, Avinashilingam University, Tamil Nadu, India, meenapriya05@gmail.com
[70] Associate Professor, Avinashilingam University, Tamil Nadu, India, shanthinelson@gmail.com

EXPERIMENTAL PROCEDURE

I. Sample Preparation

In this part, Barrium titanate thin films were preparation is discussed. Barrium titanate thin films were deposited on glass substrates. Choice of substrates posses much influence in determining thin film properties. Glass plates fulfill varied requirements of substrates best and they can be rapidly coated by spin coating technique.

The barium titanate thin films were deposited onto highly clean glass substrates. Barium titanate with the average molecular weight of 233.192 g/mol (nano powder cubic crystal phase with particle size of >100nm) and HNO3 were used for preparing the precursor solution. The solution was prepared by dissolving the known quantities of barium titanate in 50ml HNO_3 at room temperature to yield different molar solutions.

The solution was continuously stirred for about 2 hrs at 55°C temperature by means of Teflon coated magnetic stirrer to ensure homogeneous mixing of the solvent. Thus, prepared solution was sprayed onto the substrate kept on a substrate base plate of the spin coater using a nozzle. The deposition was allowed to take place for the rotation of 2000 rpm at a pressure of 20 kpas for 45 and 60 seconds respectively.

- Sample I: 1Mole of $BaTiO_3$ dissolved in HNO_3
- Sample II: 0.5Mole of $BaTiO_3$ dissolved in HNO_3

After the film deposition, coated substrates were taken out and annealed in muffle furnace at 600°C for 2 hours to remove all the traces of solvent.

RESULTS AND DISCUSSIONS

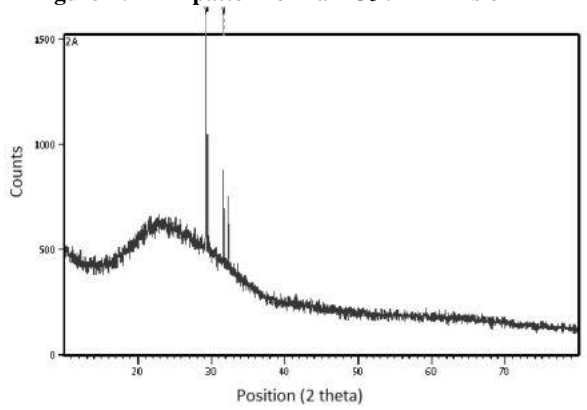
I. Film Thickness

Thickness of film is among the first quoted attribute of its nature. The reason is that thin film properties usually depend on thickness and it plays an important role in the film properties. In our present work thickness of the films were measured by adapting mechanical technique, viz microbalance gravimetric technique. This method depends on the increase of weight of a film due to its mass increase and from the knowledge of its density and deposited area, film thickness (t) was evaluated from the relation, $t=W/\rho A$. In the above relation W represents the weight of the deposited film. For all the deposited film, thickness fell in few µm ranges. Thickness increases as concentration increases, which indicate the validity of the deposition technique adopted for the film. It was observed that film thickness increases as concentration increases with ~2000 rpm for different durations. [2,3]

II. Structural Analysis

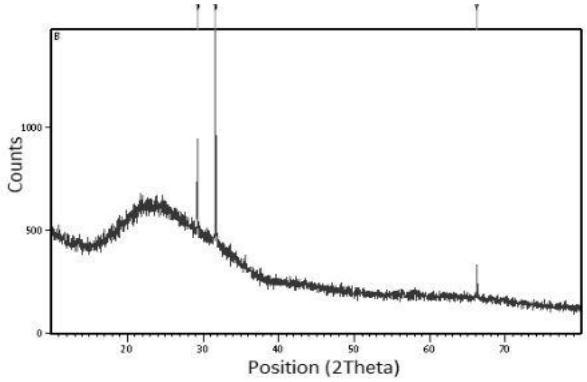
X-ray diffraction (XRD) measurements revealed films deposited on glass substrates to be essentially of [loo] orientation with trace levels of other orientations such as [llO] and [ill]. [4,7]

Figure-1: XRD pattern of $BaTiO_3$ thin films of 1M



Sources: Authors Compilation.

Figure-2: XRD pattern of $BaTiO_3$ thin films of 0.5M



Sources: Authors Compilation.

Typical XRD pattern of $BaTiO_3$ film is given in Figure 1and 2. This crystal phase structure was found by XRD method using Cu K_α radiation. The thin film sample was scanned for step size of 0.0170° for time interval of 10.1600 sec over 10.0084 to 79.9804. The sharp peak position in the spectrum, conforms crystalline nature of $BaTiO_3$.The crystallite size was estimated using Scherer formula,

$$Ps = \frac{0.9\,\lambda}{\beta\,Cos\theta}$$

Where, $\mathbf{P_s}$ is the particle size, $\boldsymbol{\lambda}$ corresponds to the wavelength of the X-ray beam, $\boldsymbol{\beta}$ is the full width of the XRD peak at half maximum and $\boldsymbol{\theta}$ is the corresponding diffraction angle. Also, inter-chain distance of the KDP doped SS film is found using the formula,

$$r = \frac{5}{8}\,\lambda\, sin\theta$$

The crystallite size and inter-chain distance with corresponding position values are shown in table 1. From the calculation of P_s from Scherer formula, it is found that the crystallite size increases with the concentration. Also, table 2 gives the values of particle size and interplanar distances for 2 different samples with varying concentrations. From table 2, it is proved that chosen $BaTiO_3$ particles fall under nano meter range. [9]

Table-1: Gives Variation of Film Thickness with Concentration.

Sample	Molar (%)	rpm	Duration (sec)	Thickness (µm)
I	1	2000	40	0.7921
II	0.5	2000	40	0.5374

Sources: Data Analysis.

Table-2: P_s and R Values

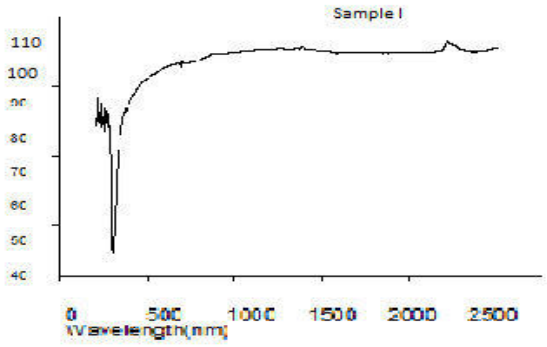
Sample	Position 2θ (rad)	P_S (nm)	R ($x10^{-13}$m)
I	0.5116	9.505	8.6083
	0.5521	9.504	9.2891
II	0.5117	9.653	8.6093
	0.5521	9.652	9.2889

Sources: Data Analysis.

III. UV-VIS Spectrum Analysis

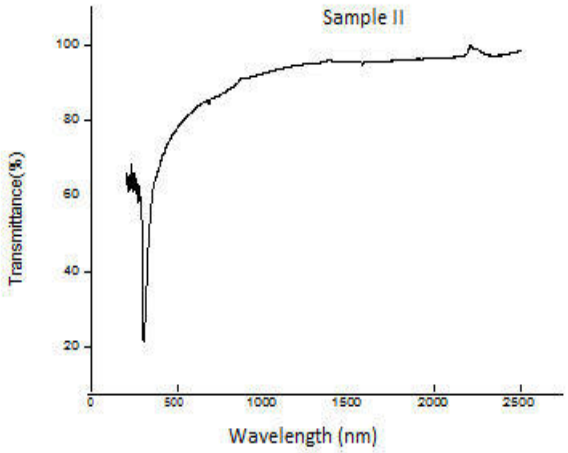
Optical studies were carried using UV-VIS spectrometer. In research, ultraviolet/visible spectroscopy is used more extensively in assaying than identification. All UV-Visible spectrometric measurements of Barrium titanate films were made with LAMBDA UV-Visible spectrometer of Perkin-Elmer. The entire coated surface was illuminated by UV-Visible radiation. The transmittance and absorbance spectra were taken in the range of 250-2500nm for thin film samples and the UV-Visible spectrums are shown in Fig3 and 4. In UV-VIS spectrum, the transmittance is defined as the ratio of the intensity of light that passes through the sample to the intensity incident upon sample. From figures 3 and 4 we observe that the zero transmittance value is around 300nm, transmittance of both the sample increases as wavelength increases, and the spectrum remains stable from 700 to 2500nm, which aids its use in electronic coatings [11].

Figure-3: UV-Vis Transmittance Spectrum of Sample I having 1M Concentration.



Sources: Authors Compilation.

Figure-4: UV-Vis Transmittance Spectrum of Sample II having 0.5 Concentration



Sources: Authors Compilation.

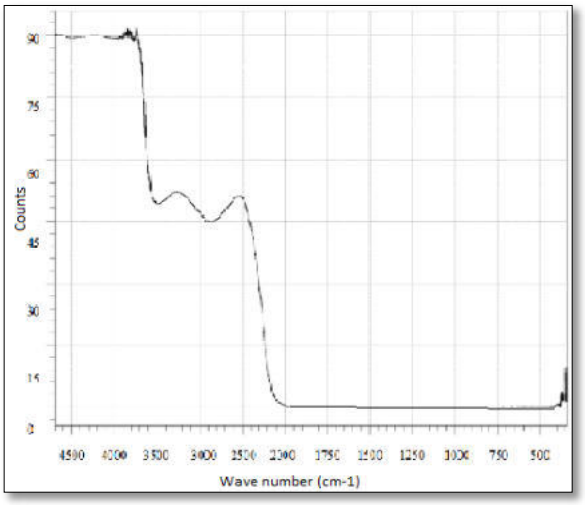
IV. FTIR Analysis

FTIR spectra recorded using Shimadzu, FTIR spectrophotometer. FTIR spectra of $BaTiO_3$ films deposited by spin coating process were recorded in wave number range of 500-4500cm^{-1} i.e. in the middle IR region, with the resolution of 4cm^{-1}.

FTIR spectrum was taken in transmittance mode and FTIR characterization spectrum of $BaTiO_3$ sample was shown in figure4. This fig shows that the vibrational peaks at 3500, 3000cm^{-1} are of pure $BaTiO_3$. [3,6]

Graph-1 shows FTIR spectrum of barium titanate thin film deposited by spin coating technique is given in which vibrational modes corresponding to our material is shown.

Graph-1



Sources: Authors Compilation.

CONCLUSIONS

In present research work, Barium Titanate ($BaTiO_3$) films have been deposited on a glass plate as it is a very attractive material in the field of electro ceramics and microelectronics due to its good characteristics.

From the UV-Visible spectrum of spin coated barium titanate the transmittance of both samples increases as wavelength increases, and spectrum remains stable from 700 to 2500nm which aids its use in electronic coatings. With the increase in concentration of solution thickness was found to increase.

REFERENCES

[1] Keigo, Suzuki, (2005). "Kazunori Kijima"; Japanese Journal of Applied Physics,44, pg. 2081-2082.

[2] M., F. Al-Kuhaili, (2006). Journal of Luminescence, 117, pp. 209–216.

[3] M., Manso-Silvana; L., Fuentes-Cobasb; R., J. MartYn-Palmaa; M., Hernandez-Veleza, and J., M. Martynez-Duarta, (2002). Surface and Coatings Technology, 151-152, pp. 118–121.

[4] Yanwei, Liu; Zhenghao, Chen; Chunling, Li; Dafu, Cui; Yueliang, Zhou, and Guozhen, Yang, (1997). Institute of Physics and Center for Condensed Matter Physics, Chinese Academy of Sciences.

[5] Sanxiong, Xiao, and Xianping, Yan, (2009). Micro Electronic Engineering, 86, pp. 387-391.

[6] Burtrand, Lee, and Jianping, Zhang, (2001). Thin Solid Films, 388, pp. 107-113.

[7] Chunlin, Fu; Wei, Cai; Longlong, Zhou; Huaqiang, Chen, and Zhirong, Liu, (2009). Applied Surface Science, 255, pp. 9444–9446.

[8] Wei, Cai; Chunlin, Fu; Jiacheng, Gao; Qian, Guo; Xiaoling, Deng; Chaoyang, Zhang, and Physica, B., (2011). 406, pp. 3583–3587.

[9] Yoshio, Kobayashi; Hirobumi, Saito; Takafumi, Kinoshita; Daisuke, Nagao, and Mikio, Konno, (2011). "Thin Solid Films", 519, pp. 1971–1975.

[10] N., G. Semaltianos, (2007). "Micro Electronics Journal", 38, pp. 754-761.

[11] Keigo, Suzuki, and Kazunori, Kijima, (2005). Japanese Journal of Applied Physics, 44, pp. 2081 - 2082.

*****

DESIGN OF WIDE RANGE MEMS TUNABLE CAPACITOR

Manoj Srinivasan R.[71] Boopathy Shanmugam P.[72]

ABSTRACT

The RF applications like voltage controlled oscillators, tunable filters, resonators etc., requires tunable capacitors in their designs. This paper presents the design of wide range MEMS tunable capacitors for RF applications. This design consists of an air suspended bottom plate and a fixed top plate. The top fixed plate and the suspended bottom plate form the tunable capacitor. The capacitance range of this tunable capacitor is from 69.172 to 138.344 nF. This range is wider compared with the conventional MEMS tunable capacitors of tuning ranges in pico Farads. The fabrication process is similar to that of the existing standard integrated circuit fabrication processes, which makes this design suitable for integrated RF applications.

KEYWORDS

MEMS Sensor, Humidity Sensor, Intellisuite, Circuits, Capacitance etc.

INTRODUCTION

MEMS owing to its higher performance and reliability in wireless and RF applications which can be used to design all the RF components in Micrometre scale (Mohamed et al. 2000a). The MEMS is an advancement of VLSI technology enables the design RF tunable components and switches. In the past few years MEMS technology is used for increasing the performance of the devices and for designing novel components (Nguyen and Howe 1999).

Nowadays, the RF tunable capacitors are designed using micromachining technology (Goldsmith et al. 1995). Apart from the conventional varactors, micro-machined tunable capacitors have better efficiency and low loss (Young et al. 1998). This paper deals with the design of MEMS-based wide range tunable capacitor structure using Intellisuite. Here we have a parallel plate capacitor with a fixed aluminium plate and a suspended movable aluminium plate for capacitance tuning. The movable plate is initially 8 µm away from the fixed plate. The maximum possible deflection of the movable plate is 4 µm towards the fixed plate. This deflection range varies the capacitance value ranging from 69.172 to 138.344 nF (Hajimiri and Lee 1999). This will be a wide tuning range in micromachined tunable capacitor structures (Wu et al. 1998).

THEORY

The basic theory behind this tunable capacitor is that the range can be increased by two ways: 1) The first method is by varying distance between the plates in larger scale (Jun Zou et al. 2001); and 2) The next method is by dynamically varying the effective area among the two plates. Here we use the first method of increasing the range of tuning, because varying the area between the plates dynamically is difficult.

[71]PG Student, Akshaya College of Engineering and Technology Tamil Nadu, India, manojece80@gmail.com
[72]PG Student, Akshaya College of Engineering and Technology Tamil Nadu, India, boopathy2702@gmail.com

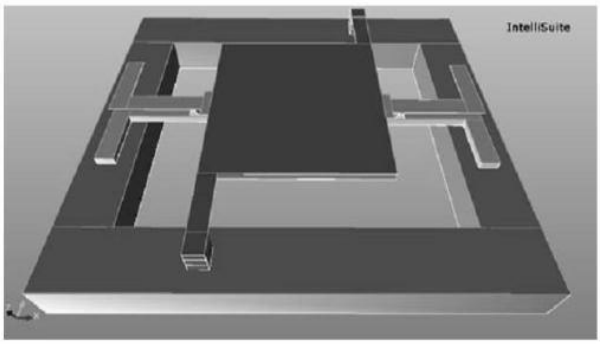
Figure-1: Initial and Final Position of Plates and Distances between Fixed and Movable Plates in Initial and Final Tuning Positions

Sources: Authors Compilation.

DESIGN

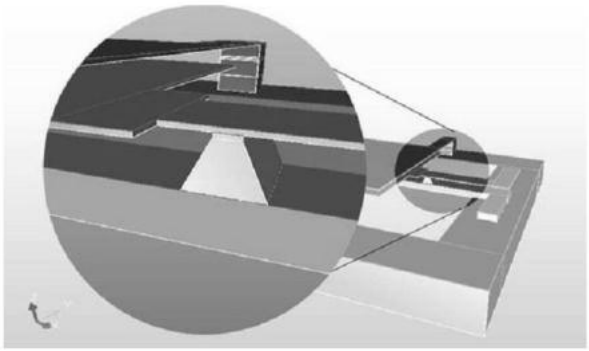
The structure of wide range MEMS tunable capacitor is shown in Fig. 2. It is designed using the 3-D builder module of the Intellisuite. Tunable capacitor structure is formed on a substrate and two aluminium plates of size 250 pm x 250 µm. Movable plate is suspended from two thermal actuators attached to the substrate (Mohamed et al. 2000b). The substrate is formed with a hole to accommodate capacitor structure over it.

Figure-2: Structure of Wide Range MEMS Tunable Capacitor



Sources: Authors Compilation.

Figure-3: Actuator-Movable Plate Contact Arrangement



Sources: Authors Compilation.

The actuator is made up of polysilicon material (Huang and Lee 1999). The actuator arrangement is given in Fig. 3. The

actuator has a small hump, which holds the plate over it. The top plate is attached with the substrate using a support above the suspended bottom plate. The distance between the plates is kept as 8 µm at the normal state. The thermal actuators have the property of thermal expansion to produce motion, when applied with a voltage. The difference in thermal expansion between two materials forms displacement. Here the actuators are applied with a voltage ranging from 0 to 20 V to get the deflection. The linear decrease in the gap between the plates can be obtained ranging from 8 to 4 µm.

ANALYSIS

The static displacement analysis of the tunable capacitor was carried out using the thermoelectromechanical module of the Intellisuite. The analysis was done on the bottom movable plate of the structure for 4 pm displacement. The deformed structure after the analysis is shown in Fig. 4.

Table-1: Theoretical Capacitance Values

Distance (µm)	Capacitance (pF)	Distance (µm)	Capacitance (pF)
24	272.813	24.6	266.159
24.1	271.781	24.7	265.082
24.2	270.559	24.8	264.013
24.3	269.445	24.9	262.953
24.4	268.341	25	261.901
24.5	267.246		

Sources: Authors Compilation.

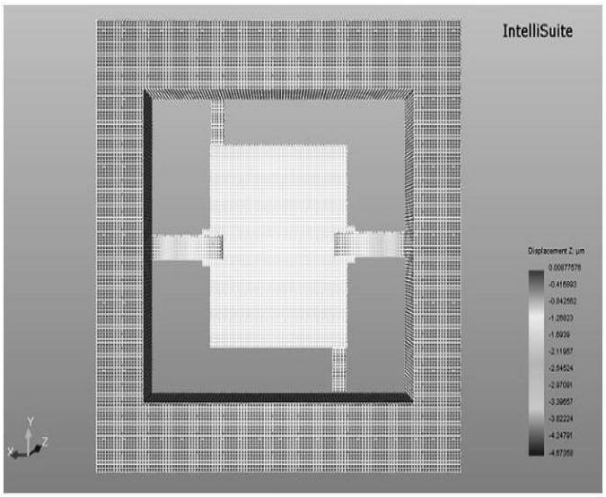
The result shows a good linear displacement of the plate on all areas. The range of the capacitance variation can be denoted as: C_{min} and $_{Cmax}$.

$$C_{min} = \varepsilon_0 * \varepsilon_r * A/D_i$$

$$C_{max} = \varepsilon_0 * \varepsilon_r * A/D_f$$

By using the above equations the capacitance range of the design is calculated. Various capacitance values for various distance between the plates is calculated and shown in Table 1. From the table it clearly shows that the capacitance is linear with the displacement with wider range of values.

Figure-4: Deformed Structure after Displacement Aalysis



Sources: Authors Compilation.

CONCLUSIONS

The wide range MEMS tunable capacitor for RF applications was designed and analysed using Intellisuite. The displacement analysis was performed on the structure and the result shows a fine and uniform displacement of the movable plate. The range of capacitance variation is calculated to be from 69.172 to 138.344 nF. This shows the design had a wide range of capacitance for RF applications.

REFERENCES

[1] Goldsmith, G.; Lin, TH; Powers, B.; Wu, WR, and Norvell, B., (1995). "Micromechanical Membrane Switches for Microwave Application", IEEE MTT-S International Symposium Digest.

[2] Hajimiri, A., and Lee, TH.., (1999). "Design Issues in CMOS Differential LC Oscillators", IEEE Journal Solid-State Circuits, 34(5), pp. 717–724.

[3] Huang, QA., and Lee, NKS., (1999). "Analytical Modeling and Optimization for a Laterally - Driven Polysilicon Thermal Actuator", Microsystems Technology, Volume: 5. Springer, New York.

[4] Mohamed, A.; Elsimary, H., and Ismail, M., (2000a). "Design of MEMS Tunable Capacitor All Metal Microstructure for RF Wireless Applications", Analog VLSI Lab, EE Department, Ohio State University, Columbus.

[5] Mohamed, A.; Elsimary, H., and Ismail, M., (2000b). "VCO Design using MEMS Technology", Proceedings of Design, Test, Integration and Packaging of MEMS/MOEMS Symposium, France (Accepted for publication).

[6] Nguyen, CT., and Howe, RT., (1999). "An Integrated CMOS Micromechanical Resonator High-Q Oscillator", IEEE Journal Solid-State Circuits, 34(4), pp. 440–455.

[7] Wu, HD.; Harsh, KF.; Irwin, RS.; Zangh, W.; Michelson, A., and Lee, C., (1998). "MEMS Design for Tunable Capacitors", IEEE MTT, Baltimore.

[8] Young, J.; Malba, V.; Ou, JJ.; Bernhardt, AF., and Boser, B., (1998). "A Low-Noise RF Voltage Controlled Oscillator using On-Chip High-Q Three-Dimensional Coil Inductor and Micromachined Variable Capacitor", IEEE Solid-State Sensor and Actuator Workshop, Hilton Head Island, SC.

[9] Zou, J.; Liu, C., and Schutt-Aine, JE., (2001). "Development of a Wide-Tuning-Range Two-Parallel-Plate Tunable Capacitor for Integrated Wireless Communication Systems", IEEE Electron Devices Meeting, 2000., IEDM Technical Digest., International, pp. 403–406. ISBN: 0-7803-6438-4.

*****

AGRICULTURAL AUTOMATION - AA

P. Srinivasan[73] S. Balamurugan[74]

ABSTRACT

The present day Indian agriculture is in downfall due to Agri land development into residential areas, lack of manpower, lack of technological up gradation, failure of seasonal rainfall, hurricane and lack of water supply. Out of all the economic activity agriculture is the prime most and should be safeguarded for the survival of mankind.

Agriculture plays an important role in life of human being. As economic activity, agriculture is a Labour intensive and involves large number of traditional processes. Farmers require simplified machineries/process to farm, to plough, to harvest and to mobilise the harvested. Next evolution in agriculture will be of Automation. Agribots should be equipped to do farming, Sowing, Irrigating, Supervising, Harvesting and Locomotion of Farm Products. In recent years the "Automation" has gained a significant role in all the areas. This paper aims to find the ways and means of achieving excellence by automation in mentioned areas to enhance field of agriculture with simple and effective techniques.

KEYWORDS

Agriculture, Automation, Methodology, Productivity, Economic etc.

NEED FOR AUTOMATION

Automation can be defined as a machine which performs the action without much of human intervention. The results of such automation are drastic reduction in labour and cost. The term automation and innovation are interrelated aims for the improvement. If there is improvement, then the below techniques should be taken for consideration. Such as Kaizen and Re-engineering were results in improvement of a process but it varies with time and cost.

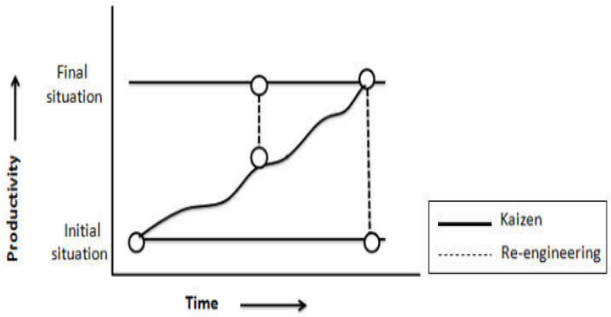
As per the figure-1, the re-engineering is the vertical dotted line, hence the productivity improves a lot in a time but with large amount of investment for example installing the robotic assembly unit in a automobile industry. Whereas kaizen involves less cost in continuous improvements than re-engineering.

For development of the agriculture, technological upgradation is done considering the above concepts in mind. Agricultural product prices are increasing day by due to increase in demand along with the reduction in supply. The reduction in supply is mainly due to labour oriented task, postpone of seasonal rainfall, issue with the sharing of water across the borders, rapid urbanisation of agriculture lands into townships, and shortage of Labours. Hence it is the time to make use of all technological upgradation in agriculture for the mankind to survive.

[73] Assistant Professor, Periyar Maniammai University, Tamil Nadu, India, charisrini@yahoo.com
[74] Assistant Professor, Periyar Maniammai University, Tamil Nadu, India, bala_guys@yahoo.co.in

KIND OF IMPROVEMENT BASED ON KAIZEN & RE-ENGINEERING

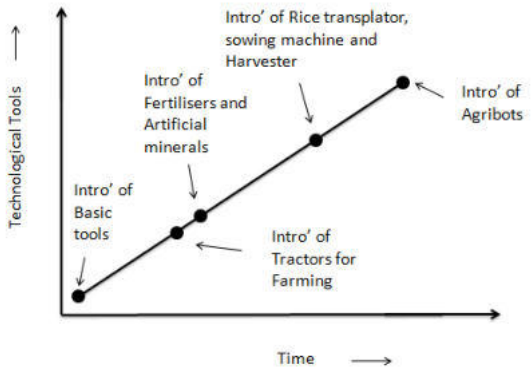
Figure-1



Sources: Authors Compilation.

INNOVATION TIME LINE IN AGRICULTURE

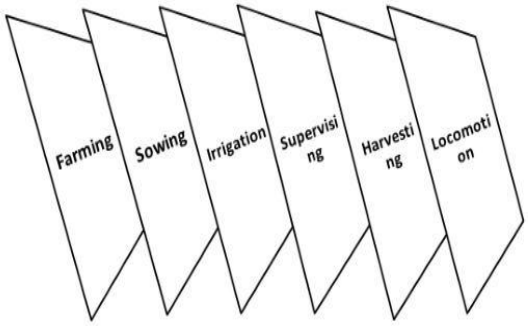
Figure-2



Sources: Authors Compilation.

FLOW CHART IN AGRICULTURE – ACTIVITIES

Figure-3



Sources: Authors Compilation.

WAYS TO AUTOMATING THE AGRICULTURE ACTIVITIES

Table-1

S.No.	Activities	Traditional Methodology	Present Methods	Proposed Technological Up gradation
1.	Farming	Wooden / Steel Plough	Farm Tractor	Automated with Sensors- AMPV*
2.	Sowing	Manual	Rice Transplator and Seed Sowing Machine	Sowing machine - AMPV
3.	Irrigating	Manual Directing Cannel / River Irrigation	Drip Irrigation System	Automatic Irrigation System
4.	Supervising	Manual Assisted with Simple Tools	Manual assisted with Weedier	Agribots for Crop monitoring and Fertilizer sprayer, Solar powered Warning System
5.	Harvesting	Manual	Combine Harvester	Compact Field Cropper - AMPV
6.	Locomotion	Manual	Manual	Multi powered bunch Transporter

Sources: Authors Compilation.

*AMPV - Agricultural Multipurpose Vehicle, a modified version of Farm Tractor to perform all major activities by a single machine through kaizen method to mechanise below activities like pumping water (in and out of the field), sowing grains, harvesting can be made by getting power from the tractor to the various attachments by this limited use tractor become AMPV.

Farming

It is the process of making the land suitable for agriculture. Here the soil gets loosened and land gets levelled by ploughing. It is done to make the root to go deeper into the soil. Traditionally it was done manually with the help of wooden / steel plougher. Presently it is replaced by the tractors for farming the land. Even in this also farmers has to depend on man for operating the tractors. The tractors also have to carry out only for ploughing and transportation with huge money investment with limited activity. This in turn not justifies our huge investment to output from the machine. It must be added with other facilities to be carry out almost all operation of agriculture

Proposed

The farm tractor is modified into AMPV, in the future AI will be applied to AMPV to become AAMPV ie., Agribots which is a fully automated robot to perform all the agriculture activities.

Sowing

It is the process of planting the seeds deep into the soil. Traditionally this is almost carried out manually. Presently this is automated with an attachment in the existing tractor. The attachment consists of a hopper with multi layer gravity feeder. By this manual process is replaced with the sowing machine for both wet and dry land. For Rice cultivation, the rice transplator is used as a separate machine exclusive for transplantation.

Proposed

For rice cultivation, the sapling is grown in nursery and transplanted using Rice Transplator. In sapling cultivation, the process of grain feeding, irrigating is fully automated. The present rice transplator is modified as attachment for AMPV.

Irrigation

It is the process of feeding water to the cultivation land or soil and draining out from the land in excess situations like rainfall cum flooding. Feeding of water is done periodically depending upon the type of crop and the condition of the land. The water level is most important one for the growth cum yield of the crops. Traditionally water is fed to cultivatable land through channels from the river or from the Deep well or from the water pump sets to the fields which have the following drawbacks. The water going through channels is wasted due to non cultivatable soil absorption (approximately 20 to 30%) The next drawback of the system is under irrigation or over irrigation .The water level management is very difficult in the existing system.

Proposed

The problem of soil absorption is greatly reduced with the introduction of Drip irrigation. For rice, Plantain and sugarcane the drip irrigation will not be effective system. The current channel irrigation is to be replaced with PVC pipe lines and automated valves to feed/stop the water in and out to the field. The water level in the field is sensed by the level/Float sensors.

Based on the feedback from the sensor the pump can be regulated. With this proposed techniques we can overcome the drawbacks of the traditional system. The moisture sensors are introduced in the field which is programmed with preset value. Whenever the field moisture is reduced from the preset value the feedback signal operates the motor equipped with controller and vice versa. The message will be send to farmer as information.

Supervising

It is a process of monitoring the field crops from insects, crop disease, animal roaming, and crop theft. Traditionally supervising is done by human patrolling.

Proposed

The spider cam technology is used for monitoring field and fertilizer spraying. Further, solar powered warning system is to be introduced. This system aims to create panic among those who (Animals, strangers) wandering near the crop field. In case of any unwanted movement wireless motion detecting sensors gives the signal to the amplifier which plays the pre-recorded voice randomly. This will create caution for the former as well as stranger or any animal which is near to the crop field

Harvesting

Traditionally the harvesting is done manually for all type of crops with the help of simple cutting tools. Presently combined harvester is used. Harvester can cut the crop from the field and it can separate the grains from the straw. Special type of harvester machines can perform packaging too. Even though the present harvester is widely accepted, there are some flaws observed by the farmers. Such as Massive structure with very huge investment to buy the tractor, grain loss during threshing (It is the process of crushing the reaped crops, so that the grains are separated from the crops) and winnowing (It is the of separation process that separates grain from the straw through sieving process), for wet land harvesting is difficult to perform, and straws wasted in the field. Finally it is difficult to afford by the farmer to buy machines for each activity.

Proposed
The AMPV is equipped with compact field cropper which is under the design of prototype, which is to be developed to overcome the flaws in the present harvester.

Locomotion

It is the process of shifting of agriculture outputs from the field to the transporting vehicle. In most cases, the crop field is not accessible by the trucks hence the shifting and loading is done manually.

Proposed
The Multi bunch transporter is introduced to carry out the locomotion which is handy to carry and install. It contains a motor with belt driven that can be easily mounted in a tractor and the support is in the farm land. If the farmers require for a larger distance it can be extended.

ACTION PLAN TO OVERCOME THE DIFFICULTIES IN IMPLEMENTATION

In spite of all such technological evolution still there is some difficulties in implementation as like the below mentioned and how we can overcome the same, are also suggested here:

1. **Financial Problem**
   Investing can be made easier with the subsidies given to farmer's group and purchasing of vehicle by small group farmer Society. This in turn breaks single individual's huge investment and sharing of high end technology.

2. **Governmental Support**
   Currently Government is giving subsidies to purchase imported Agri equipments, Funding cum tax rebate support to be made for Indian companies to manufacture high end agricultural machineries & further Agricultural research works for economy and growth.

3. **Lack of Technological Capability-cum-Initiative**
   Government, public and private sector is to take initiative for agricultural automation product development and collaborative works with reputed organisations of abroad to bring the latest innovations to Indian agriculture.

4. **Lack of Awareness**
   The farmers must be exposed to recent innovations by giving training programmes / videos and also make initiatives in hands on training of automated equipments.

CONCLUSIONS

The paper gives the broad spectrum of the current trend of Indian agriculture. The main aim is to create the awareness cum current need about agricultural automation and the scope of opportunities available with Indian agricultural field. This in turn ushers our industrials' to turn cum focus their vision into agricultural automation related product development as like our automotive sector. As part of government also to fully extend their support by giving subsidies, tax rebate and support in collaborative works to the agricultural automation product manufactures. This creates our future generation from food crisis cum starvation, countries economy growth, higher productivity, Reduces unemployment related problems.

REFERENCES

[1] R., Ceres; J., L. Pons; A., R. Jiménez; J., M. Martín, and L., Calderón, (1998). "Design and Implementation of an Aided Fruit-Harvesting Robot (Agribot)", Industrial Robot: An International Journal, Volume: 25 Issue: 5, pp.337–346.

[2] K., E. Bouazza; M., Ouali, and M., Boutayeb, (2009). "Modelling, Identification and Control of Three Axis Manipulator Robot", Journal of Automation & Systems Engineering, Volume: 3, Issue: 4, pp. 2.

[3] J., K. Schueller, (2006). Automation and Control. In: CIGR Handbook of Agricultural Engineering, Information Technology, Volume: VI, ed., by A. Munack (CIGR, Tzukuba 2006), pp. 184–195, Chapter: 4.

[4] Theodore, J. Williams, (2009). "Advances in Industrial Automation: Historical Perspectives", Springer Handbook of Automation, pp. 5-11.

[5] J., Feyen, and F., Liu, (1991). "Automation of Design of Agricultural Water Management Projects", Springer: Water Resources Management, Volume: 5, Issue: 2, pp. 95-119.

[6] Ravi, Anupindi, and S., Sivakumar, (2007). "Supply Chain Reengineering in Agri-Business", Springer: International Series in Operations Research & Management Science, Building Supply Chain Excellence in Emerging Economies, Part-III, Volume: 98, pp. 265-307.

[7] Imai, Masaaki (1986). Kaizen: The Key to Japan's Competitive Success, New York, NY, US: Random House.

*****

AERODYNAMIC CHARACTERISTICS OF MODIFIED BLUNT TRAILING EDGE AIRFOIL

I. Balaguru[75] R. Rajesh[76] M. Manikandan[77]
K. Sampathraja[78] S. Subash[79]

ABSTRACT

This paper discusses the aerodynamic performance of a modified blunt trailing edge airfoil has been analysed using a two-dimensional Unsteady Reynolds Averaged Navier Stokes (URANS) method. The blunt edge airfoils usually used for the compressor blades, wind turbine blades, etc because of its large thickness. The drag is much more when compared to the conventional airfoils, because of its wake formation near the ends. The drag is the imperative criteria which must be reduced for the effective performance of the aircraft during take-off conditions.

The usage of mechanism in aircraft wing structures reduces the drag to some extent, but it has to be reduced further. To overcome this, the trailing edge of the blunt airfoils has been modified with V shape at its blunt end. The wake near the trailing edge of the airfoils is also reduced in the blunt airfoils. The conventional and blunt trailing edge airfoil has been modelled and the numerical analysis is carried out using commercial numerical codes to study its aerodynamic performance. The results describe the changes in lift and drag characteristics and changes in stall angle of attack.

KEYWORDS

Aerodynamic, Unsteady Reynolds Averaged Navier Stokes (URANS), Airfoil, CFD, Blunt Trailing Edge Airfoil (BTEA) etc.

INTRODUCTION

Conventionally, airfoil geometries are such that they have a smoothed leading edge and a sharp trailing edge. It would be useful in some cases, to use airfoil geometries which are thicker, in order to attain better structural performance. The effects on aerodynamic performance caused by modifications made to the trailing edge of subsonic profiles used in air vehicles.

The blunt edge airfoils are reauired in various aerodynamic systems, including blended wing-body aircraft, unmanned aerial vehicles, and wind turbine blades, while reducing the well-documented separation sensitivity of airfoils with maximum thickness-tochord ratios greater than 25%.

Blunt trailing edge airfoils provide several structural and aerodynamic performance advantage rather than conventional airfoils [1]. Structurally, [2] this airfoil increases the sectional area and sectional moment of inertia for a given airfoil maximum thickness Aerodynamically, it increases sectional maximum lift coefficient and lift curve slope, and reduces well-documented sensitivity of lift characteristics of thick air-foils.

Incorporation of trailing-edge thickness allows for a portion of the pressure recovery to occur in the airfoil wake, thereby reducing the severity of the adverse pressure gradient on the suction surface. This reduction in the adverse pressure gradient on the suction surface alleviates the tendency toward premature boundary layer separation for both clean and soiled conditions and improves lift performance [3].

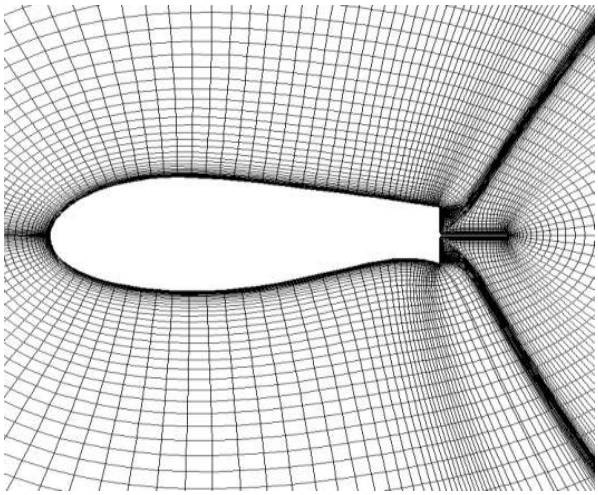
From the past research works, it can be concluded that the lift to drag ratio is decreased in blunt trailing edge airfoils due to the thickness at its end. To overcome this problem, researches used some mechanisms which includes gurney-type tabs, splitter plates, base cavities, offset cavities, etc, to reduce drag.

In this research work, the aerodynamic performance of the blunt edge trailling edge airfoils are improved by making 'V' shaped changes at the blunt end. URANS based numerical investigations are carried out to predict aerodynamic performance of the modified blunt airfoils.

BLUNT AIRFOIL

Blunt trailing-edge [5] airfoils have been proposed to address the need for thick airfoil sections necessary to meet structural and volume requirements of various aerodynamic systems, including blended wing-body aircraft, unmanned aerial vehicles, and wind turbine blades, while reducing the well-documented separation sensitivity of airfoils with maximum thickness-tochord ratios greater than 25%. Without some sort of separation control technique, such as vortex generators, boundary layer suction, or a blunt trailing edge, thick airfoils generate steep adverse pressure gradients, which lead to premature flow separation from upper surface with even small perturbations.

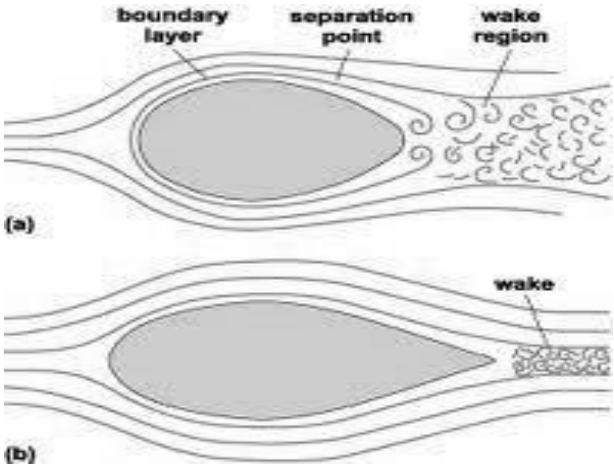
Figure-1: Blunt Airfoil with Splitter Plate



Sources: Authors Compilation.

[75]Assistant Professor, Karpagam Institute of Technology, Tamil Nadu, India, ibalaguru@yahoo.co.in
[76]UG Student, Karpagam Institute of Technology, Tamil Nadu, India, rajesh657310@gmail.com
[77]UG Student, Karpagam Institute of Technology, Tamil Nadu, India, manimurugesan@gmail.com
[78]UG Student, Karpagam Institute of Technology, Tamil Nadu, India, sampathraja92@gmail.com
[79]UG Student, Karpagam Institute of Technology, Tamil Nadu, India, subash.mail14@gmail.com

Figure-2: Wake Formation in Aerofoil



Sources: Authors Compilation.

Several studies have investigated these so-called "thick" airfoils and show that blunt trailing-edge airfoils have significantly improved lift performance compared to similar sharp trailing-edge airfoils.

The use of blunt trailing-edge, however, introduces base drag and possible vortex shedding, which in most cases is undesireable.In order to improve the performance characteristics of blunt trailing-edge airfoils, some form of drag mitigation is necessary.

Past studies have investigated several options for bluff body drag reduction, but most of this research is directed toward two-dimensional axisymmetric bodies aligned with flowfield. Limited research has been conducted for bodies at incidence to the flow and even less study has been conducted for asymmetric lifting bodies, such as blunt trailing-edge airfoils.

In the present study, the results of a coupled computational fluid dynamics (CFD) and wind tunnel experimentation study are presented for simple, static, trailing-edge attachments applied to a blunt trailing-edge airfoil in two-dimensional flow.

WAKE

A wake is the region of recirculating flow immediately behind a moving or stationary solid body, caused by the flow of surrounding fluid around the body which is shown in the fig.2.

In fluid dynamics, [4] a wake is the region of disturbed flow (usually turbulent) downstream of a solid body moving through a fluid, caused by the flow of the fluid around the body. In incompressible fluids (liquids) such as water, a bow wake is created when a watercraft moves through the medium; as the medium cannot be compressed, it must be displaced instead, resulting in a wave.

As with all wave forms, it spreads outward from the source until its energy is overcome or lost, usually by friction or dispersion. The formation of these waves in liquids is analogous to the generation of shockwaves in compressible flow, such as those generated by rockets and aircraft traveling supersonic through air [5].

DESIGN

The geomerty of blunded trailing edge airfoils are desinged with maximum of 21% of thickness to chord ratio and the trailing edge is desinged in curved shape, which is used to reduce more amount of drag. The geometry has been desinged in effective manner to produce better performance rather than the conventional airfoils.The airfoil choosen for making blunded airfoil is NACA 0021, because of its simple geometry.

In this prposed research work, the NACA 0021 airfoil's trailing edge is modified in to V-shape. Due to the modified design, the pressure difference is high without the usage of splitterplate mechanisms. Thus the performance of the NACA 0021 conventional airfoil & modified blunted aitfoil can be compared.

RESULTS & DISCUSSIONS

In-order to predict the aerodynamic characterisics of the modified blunt edge airfoils, the numerical analysis is carried out at 70m/s. The aerodynamic characteristics such as pressure distribution, velocity distriution, lift and drag are for the conventional and modified blunt airfoils are discussed.

Pressure Distribution

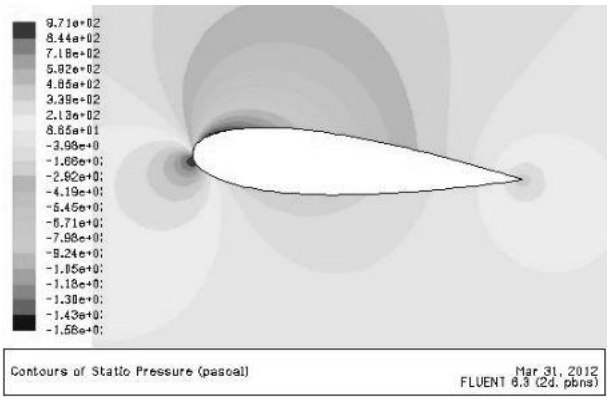
Pressure distribution plays a vital role in aircraft aerodynamic performance analysis.The pressure distribution of conventional and modified blunt trailing edge airfoil has been discussed. The result predicted for angle of attack in range of 0°, 5°, 10°, 15° for calculate the aerodynamic characteristics of airfoils.

Fig.3a and Fig.3b illustrates pressure distribution over conventional and modified blunt airfoil of same cross section at 5° angle of attack.

Fig.4a and fig.4b illustrates pressure distribution over conventional and modified blunt airfoil of same cross section at 5° angle of attack. From the result, it is seen that the pressure of blunt airfoil on the top surface pressure is very low in comparison with conventional airfoils.

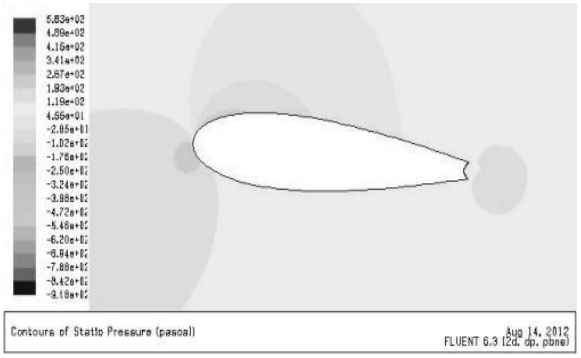
Moreover, the pressure difference in modified blunt airfoil is high at the 15° angle of attack. Pressure distribution indicates the effective performance of that blunt airfoil.

Figure-3 (a): Convemtional Aerofoil, Pressure Distribution for 5° Angle of Attack



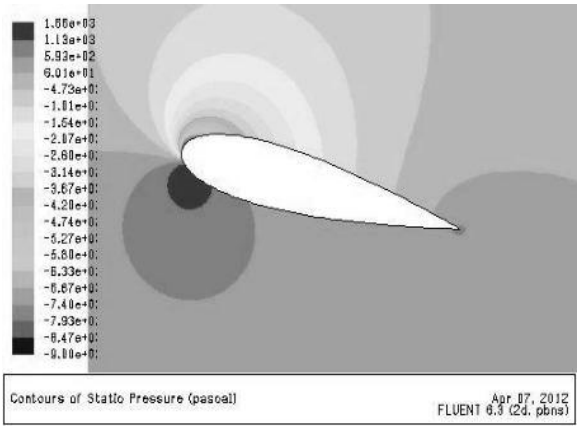
Sources: Authors Compilation.

Figure-3 (b): Modified Blunted Aerofoil, Pressure Distribution for 5° Angle of Attack



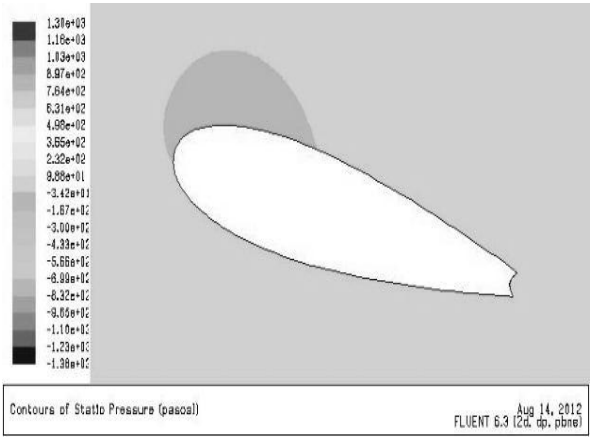
Sources: Authors Compilation.

Figure-4 (a): Convemtional Aerofoil, Pressure Distribution for 15° Angle of Attack



Sources: Authors Compilation.

Figure-4 (b): Modified Blunted Aerofoil, Pressure Distribution for 15° Angle of Attack

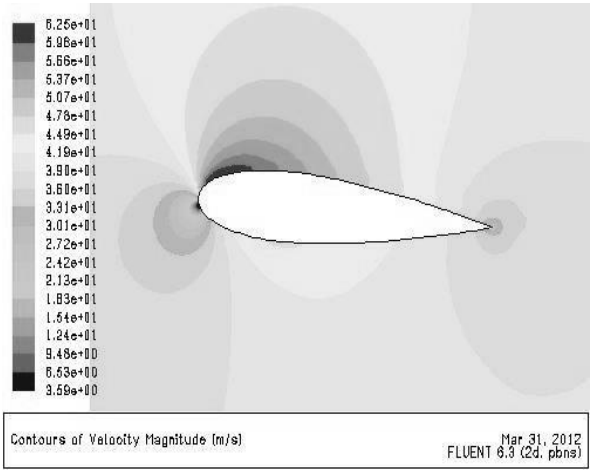


Sources: Authors Compilation.

Velocity Distribution

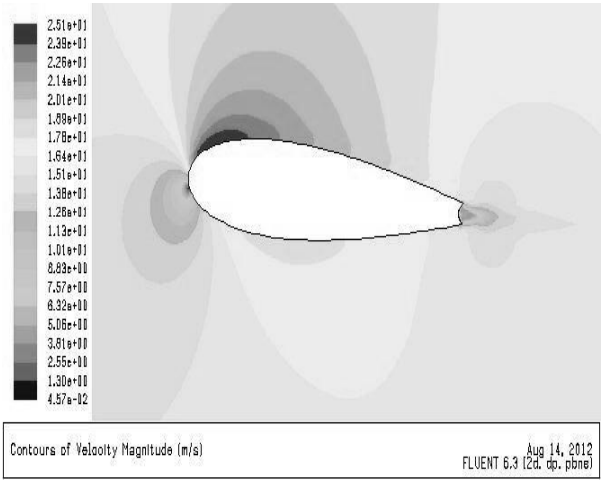
Fig 5, 6 shows velocity distribution for conventional and blunt airfoil at 5° and 15°. It is clearly seen that the velocity distribution over the surface of blunt wing is even though low, the lift of the blunt wing is very high, and the values of drag is too low. The separation of boundary layer is reduced if the modified blunt airfoil in used.

Figure-5 (a): Convemtional Aerofoil, Velocity Distribution for 5° Angle of Attack



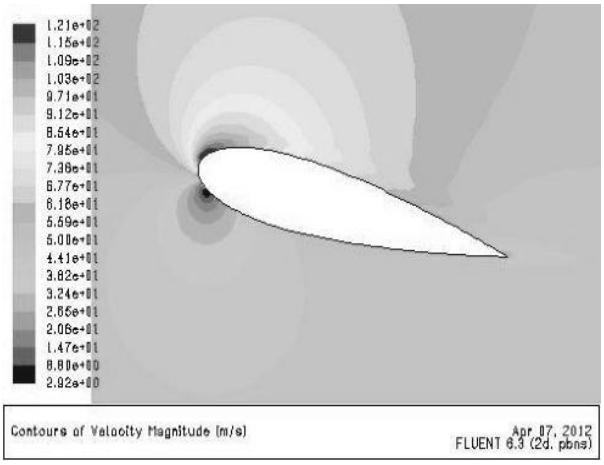
Sources: Authors Compilation.

Figure-5 (b): Blunted Aerofoil, Velocity Distribution for 5° Angle of Attack



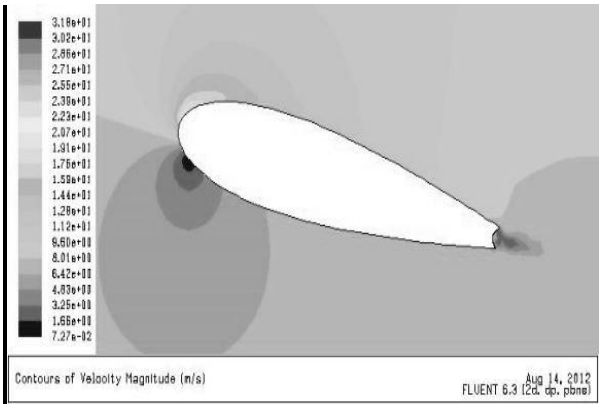
Sources: Authors Compilation.

Figure-6 (a): Convemtional Aerofoil, Velocity Distribution for 15° Angle of Attack



Sources: Authors Compilation.

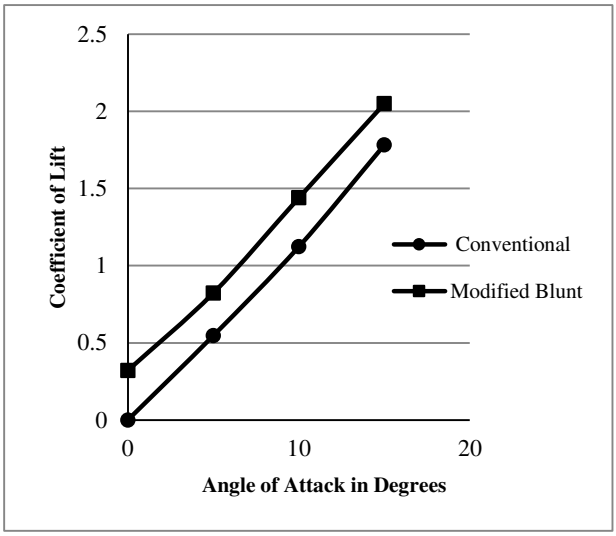
Figure-6 (b): Blunted Aerofoil, Velocity Distribution for 15° Angle of Attack



Sources: Authors Compilation.

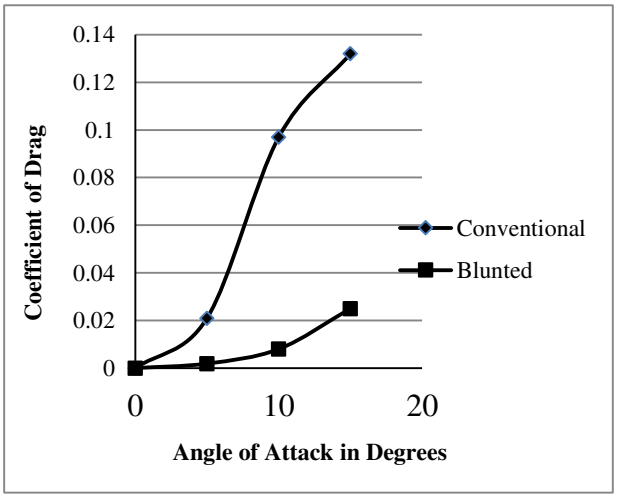
LIFT AND DRAG

Figure-7: Comparision of C_L for Various Angle of Attack



Sources: Authors Analysis.

Figure-8: Comparision of CD for Various Angle of Attack



Sources: Authors Analysis.

From the figures (7 & 8), it is clear that can get that the C_L (Coefficient of Lift) of modified blunt end airfoil is increased by 30%, thus it is obivious that lift produced by blunted airfoil with V-cut at the trailing edge will be high,when compared wirh the conventional airfoil.

CONCLUSIONS

The design and analysis of conventional and blunt airfoil has been done. The effective performance of the both airfoil has been studied. From the analysis it is clear that the drag of modified blunt airfoils is very much reduced in comparison with conventional airfoils of same cross section. In blunted airfoil by using spillter plates, wake formation will be produced which is very useful to produce lift.

In this investigation, the results are predicted by cutting the blunted airfoil trailing edge by V-shape without using splitter plates. By this process, wake formation will be there more, while compared with that of the blunted airfoil with spillter plates. Thus as a result, the maximum lift can be achieved.

REFERENCES

[1] Con, Doolan, (2007). "Numerical Simulation of a Blunt Airfoil wake using a Two-Dimensional URANS Approach", March.

[2] C., J. Doolan, (2007). "Numerical Simulation of a Blunt Airfoil Wake", 16th Australasian Fluid Mechanics Conference Crown Plaza, Gold Coast, Australia, 2nd – 7th December.

[3] Yunus, A. Cengel, and John, M. Cimbala, (2009). "Fluid Mechanics - Fundamentals and Applications", 6th Edition, Tata McGraw Hill.

[4] L., N. Clancy, (2006). The Fundamental of Aerodynamics: Sterling Book House.

[5] Jonathon, P. Baker, and C., P. Van Dam, (2008). "Drag Reduction of Blunt Trailing – Edge Airfoils", BBAA VI International Colloquium on: Bluff Bodies Aerodynamics & Applications Milano, Italy, July, pp. 20–24.

[6] John, D. Anderson, "Fundamentals of Aerodynamics", 4th Edition, McGraw Hill Higher Education.

*****

BLOCKING THE ILLEGITIMATE USERS IN ANONYMITY

Sheeba K.[80] Jeditiah R.[81] Narmadha C.[82]

ABSTRACT

Anonymizing networks such as Tor allow users to access Internet services privately by using a series of routers to hide the client's IP address from the server. The success of such networks, however, has been limited by users employing this anonymity for abusive purposes such as defacing popular Web sites. Web site administrators routinely rely on IP-address blocking for disabling access to misbehaving users, but blocking IP addresses is not practical if the abuser routes through an anonymizing network.

As a result, administrators block all known exit nodes of anonymizing networks, denying anonymous access to misbehaving and behaving users alike. To address this problem, we present Nymble, a system in which servers can "blacklist" misbehaving users, thereby blocking users without compromising their anonymity.

Our system is thus agnostic to different servers' definitions of misbehavior — servers can blacklist users for whatever reason, and the privacy of blacklisted users is maintained.

KEYWORDS

Network, Websites, Administrator, IP, Verifier-local revocation (VLR), Nymble Manager (NM), PseudoNym Manager (PNM), Back Listing (BL) etc.

I. INTRODUCTION

ANONYMIZING networks such as Tor route traffic through independent nodes in separate administrative domains to hide a client's IP address. Unfortunately, some users have misused such networks—under the cover of anonymity, users have repeatedly defaced popular Web sites such as Wikipedia. Since Web site administrators cannot blacklist individual malicious users' IP addresses, they blacklist the entire anonymizing network. Such measures eliminate malicious activity through anonymizing networks at the cost of denying anonymous access to behaving users. In other words, a few "bad apples" can spoil the fun for all. There are several solutions to this problem, each providing some degree of accountability. In pseudonymous credential systems users log into Web sites using pseudonyms, which can be added to a blacklist if a user misbehave.

Unfortunately, this approach results in pseudonymity for all users, and weakens the anonymity provided by the anonymizing network.Anonymous credential systems employ group signatures. Basic group signatures allow servers to revoke a misbehaving user's anonymity by complaining to a group manager. Servers must query the group manager for every authentication, and thus, lacks scalability. Traceable signatures allow the group manager to release a trapdoor that allows all signatures generated by a particular user to be traced; such an approach does not provide the backward unlinkability that we desire, where a user's accesses before the complaint remain anonymous. Backward unlinkability allows for what we call subjective blacklisting, where servers can blacklist users for whatever reason since the privacy of the blacklisted user is not at risk. In contrast, approaches without backward unlinkability need to pay careful attention to when and why a user must have all their connections linked, and users must worry about whether their behaviors will be judged fairly.

Subjective blacklisting is also better suited to servers such as Wikipedia, where misbehaviors such as questionable edits to a Webpage, are hard to define in mathematical terms. In some systems, misbehavior can indeed be defined precisely. For instance, double spending of an "e-coin" is considered an mis behavior in anonymous e-cash systems following which the offending user is deanonymized.Unfortunately, such systems work for only narrow definitions of misbehavior—it is difficult to map more complex notions of misbehavior onto "double spending" or related approaches.With dynamic accumulators a revocation operation results in a new accumulator and public parameters for the group, and all other existing users' credentials must be updated, making it impractical.

Verifier-local revocation (VLR) fixes this shortcoming by requiring the server ("verifier") to perform only local updates during revocation. Unfortunately, VLR requires heavy computation at the server that is linear in the size of the blacklist. For example, for a blacklist with 1,000 entries, each authentication would take tens of seconds, a prohibitive cost in practice. In contrast, our scheme takes the server about one millisecond per authentication, which is several thousand times faster than VLR. We believe these low overheads will incentivize servers to adopt such a solution when weighed against the potential benefits of anonymous publishing (e.g., whistle-blowing, reporting, anonymous tip lines, activism, and so on.).

Our Solution

We present a secure system called Nymble, which provides all the following properties: anonymous authentiction, backward unlinkability, subjective blacklisting, fast authentication speeds, rate-limited anonymous connections, revocation auditability (where users can verify whether they have been blacklisted), and also addresses the Sybil attack to make its deployment practical. In Nymble, users acquire an ordered collection of nymbles, a special type of pseudonym, to connect toWebsites. Without additional information, these nymbles are computationally hard and hence, using the stream of nymbles simulates anonymous access to services. Web sites, however, can blacklist users by obtaining a seed for a particular nymble, allowing them to link future nymbles from the same user—those used before the complaints remain unlinkable.

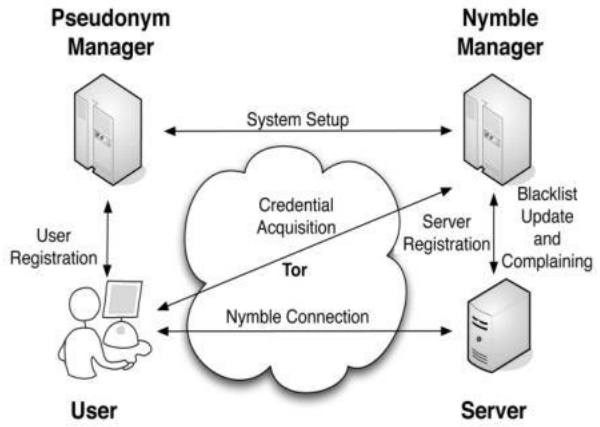
Servers can therefore blacklist anonymous users without knowledge of their IP addresses while allowing behaving users to connect anonymously. Our system ensures that users are aware of their blacklist status before they present a nymble, and disconnect immediately if they are blacklisted. Although our work applies to anonymizing networks in general, we consider

[80]Student, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, ksheeba20@gmail.com
[81]Student, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, jeditiah@gmail.com
[82] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, narmuchandran@gmail.com

Tor for purposes of exposition. In fact, any number of anonymizing networks can rely on the same Nymble system, blacklisting anonymous users regardless of their anonymizing network(s) of choice.

II. NYMBLESYSTEM ARCHITECTURE

Figure-1



Sources: Authors Compilation.

Resource-Based Blocking

To limit the number of identities a user can obtain (called the Sybil attack, the Nymble system binds nymbles to resources that are sufficiently difficult to obtain in great numbers. For example, we have used IP addresses as the resource in our implementation, but our scheme generalizes to other resources such as email addresses, identity certificates, and trusted hardware. We address the practical issues related with resource-based blocking in, and suggest other alternatives for resources. We do not claim to solve the Sybil attack. This problem is faced by any credential system and we suggest some promising approaches based on resource-based blocking since we aim to create a real-world deployment.

The Pseudonym Manager

The user must first contact the Pseudonym Manager (PM) and demonstrate control over a resource; for IP-address blocking, the user must connect to the PM directly (i.e., not through a known anonymizing network), as shown in Fig. 1. We assume the PM has knowledge about Tor routers, for example, and can ensure that users are communicating with it directly.Pseudonyms are deterministically chosen based on the controlled resource, ensuring that the same pseudonym is always issued for the same resource.

Note that the user does not disclose what server he or she intends to connect to and the PM's duties are limited to mapping IP addresses (or other resources) to pseudonyms. As we will explain, the user contacts the PM only once per linkability window (e.g., once a day).

The Nymble Manager

After obtaining a pseudonym from the PM, the user connects to the Nymble Manager (NM) through the Nymble system architecture showing the various modes of interaction. Note that users interact with the NM and servers though the anonymizing network.

Two nymbles are linked if one can infer that they belong to the same user with probability better than random guessing.Note that if a user connects through an unknown anonymizing network or proxy, the security of our system is no worse than that provided by real IP-address blocking, where the user could have used an anonymizing network unknown to the server.

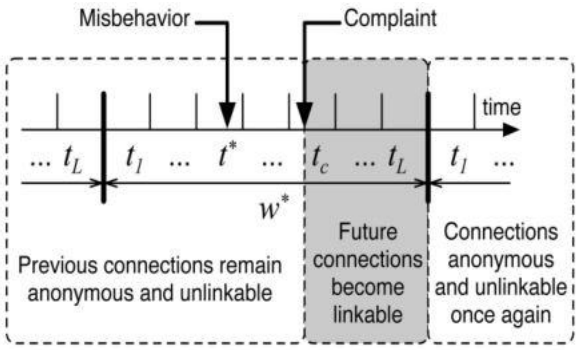
Anonymizing network, and requests nymbles for access to a particular server (such as Wikipedia). A user's requests to the NM are therefore pseudonymous, and nymbles are generated using the user's pseudonym and the server's identity. These nymbles are thus specific to a particular user-server pair. Nevertheless, as long as the PM and the NM do not collude, the Nymble system cannot identify which user is connecting to what server; the NM knows only the pseudonym-server pair, and the PM knows only the user identity-pseudonym pair.

To provide the requisite cryptographic protection and security properties, the NM encapsulates nymbles within nymble tickets. Servers wrap seeds into linking tokens, and therefore, we will speak of linking tokens being used to link future nymble tickets. The importance of these constructs will become apparent as we proceed.

Time

Nymble tickets are bound to specific time periods. As illustrated in Fig. 2, time is divided into linkability windows of duration W, each of which is split into L time periods of duration T (i.e., W =L *T). We will refer to time periods and linkability windows chronologically as t1,t2, . . . t and w1, w2, . . . , respectively. While a user's access within a time period is tied to a single nymble ticket, the use of different nymble tickets across time periods grants the user anonymity between time periods. Smaller time periods provide users with higher rates of anonymous authentication, while longer time periods allow servers to rate-limit the number of misbehaviors from a particular user before he or she is blocked. For example, T could be set to five minutes, andW to one day (and thus, L =288). The linkability window allows for dynamism since resources such as IP addresses can get reassigned and it is undesirable to blacklist such resources indefinitely, and it ensures forgiveness of misbehavior after a certain period of time. We assume all entities are time synchronized (for example, with time.nist.gov via the Network Time Protocol (NTP)), and can thus calculate the current linkability window and time period.

Figure-2: The Life Cycle of a Misbehaving User



Sources: Authors Compilation.

If the server complains in time period tc about a user's connection in t_, the user becomes linkable starting in tc. The complaint in tc can include nymble tickets from only tc_1 and earlier.

Blacklisting A User

If a user misbehaves, the server may link any future connection from this user within the current linkability window (e.g., the same day). Consider Fig. 2 as an example: A user connects and misbehaves at a server during time Period t*within linkability window w*. The server later detects this misbehavior and complains to the NM in time period tc (tc (t*< tc <=tl) of the same linkability window w*.

As part of the complaint, the server presents the nymble ticket of the misbehaving user and obtains the corresponding seed from the NM. The server is then able to link future connections by the user in time periods tc, tc1, ... tl of the same linkability window w* to the complaint. Therefore, once the server has complained about a user, that user is blacklisted for the rest of the day, for example (the linkability window). Note that the user's connections in t1, t2 . . . t*,t*+1,….tc remain unlinkable (i.e., including those since the misbehavior and until the time of complaint).

Even though misbehaving users can be blocked from making connections in the future, the users' past connections remain unlinkable, thus providing backward unlinkability and subjective blacklisting.

Notifying User of Blacklist Status

Users who make use of anonymizing networks expect their connections to be anonymous. If a server obtains a seed for that user, however, it can link that user's subsequent connections. It is of utmost importance then that users be notified of their blacklist status before they present a nymble ticket to a server. In our system, the user can download the server's blacklist and verify her status.

If blacklisted, the user disconnects immediately. Since the blacklist is cryptographically signed by the NM, the authenticity of the blacklist is easily verified if the blacklist was updated in the current time period (only one update to the blacklist per time period is allowed). If the blacklist has not been updated in the current time period, theNMprovides servers with "daisies" every time period so that users can verify the freshness of the blacklist ("blacklist from time period told is fresh as of time period tnow").

These daisies are elements of a hash chain, and provide a lightweight alternative to digital signatures. Using digital signatures and daisies, we thus ensure that race conditions are not possible in verifying the freshness of a blacklist.Auser is guaranteed that he or she will not be linked if the user verifies the integrity and freshness of the blacklist.

III. SECURITY MODEL

Nymble aims for four security goals. We provide informal definitions here, a detailed formalism can be found in our technical report, which explains how these goals must also resist coalition attacks.

Security Goals and Threats

An entity is honest when its operations abide by the system's specification. An honest entity can be curious: it attempts to infer knowledge from its own information (e.g., its secrets, state, and protocol communications).

An honest entity becomes corrupt when it is compromised by an attacker, and hence, reveals its information at the time of compromise, and operates under the attacker's full control, possibly deviating from the specification. Blacklistability assures that any honest server can indeed block misbehaving users. Specifically, if an honest server complains about a user that misbehaved in the current linkability window, the complaint will be successful and the user will not be able to "nymble-connect," i.e., establish a Nymble-authenticated connection, to the server successfully in subsequent time periods (following the time of complaint) of that linkability window.

Rate-limiting assures any honest server that no user can successfully nymble-connect to it more than once within any single time period.

Nonframeability guarantees that any honest user who is legitimate according to an honest server can nymble-connect to that server. This prevents an attacker from framing a legitimate honest user, e.g., by getting the user blacklisted for someone else's misbehavior. This property assumes each user has a single unique identity. When IP addresses are used as the identity, it is possible for a user to "frame" an honest user who later obtains the same IP address. Nonframeability holds true only against attackers with different identities (IP addresses).

A user is legitimate according to a server if she has not been blacklisted by the server, and has not exceeded the rate limit of establishing Nymble connections. Honest servers must be able to differentiate between legitimate and illegitimate users.

Anonymity protects the anonymity of honest users, regardless of their legitimacy according to the (possibly corrupt) server; the server cannot learn any more information beyond whether the user behind (an attempt to make) a nymble connection is legitimate or illegitimate.

Trust Assumptions

Table-1: Who Trusts? Whom To Be? How For? What Guarantee?

Who	Whom	How	What
Servers	PM & NM	honest	*Blacklistability & Rate-limiting*
Users	PM & NM	honest	*Non-frameability*
Users	PM	honest	*Anonymity*
Users	NM	honest & not curious	*Anonymity*
Users	PM or NM	honest	*Non-identification*

Sources: Authors Compilation.

We allow the servers and the users to be corrupt and controlled by an attacker. Not trusting these entities is important because encountering a corrupt server and/or user is a realistic threat. Nymble must still attain its goals under such circumstances.

With regard to the PM and NM, Nymble makes several assumptions on who trusts whom to be how for what guarantee. We summarize these trust assumptions as a matrix in Fig. 3. Should a trust assumption become invalid, Nymble will not be able to provide the corresponding guarantee.

For example, a corrupt PM or NM can violate Blacklistability by issuing different pseudonyms or credentials to blacklisted users. A dishonest PM (resp., NM) can frame a user by issuing her the pseudonym (resp., credential) of another user who has already been blacklisted.

To undermine the Anonymity of a user, a dishonest PM (resp., NM) can first impersonate the user by cloning her pseudonym (resp., credential) and then attempt to authenticate to a server—a successful attempt reveals that the user has already made a connection to the server during the time period.

Moreover, by studying the complaint log, a curious NM can deduce that a user has connected more than once if she has been complained about two or more times. As already described, the user must trust that at least the NM or PM is honest to keep the user and server identity pair private.

IV. CONCLUSIONS

We have proposed and built a comprehensive credential system called Nymble, which can be used to add a layer of accountability to any publicly known anonymizing network.

Servers can blacklist misbehaving users while maintaining their privacy, and we show how these properties can be attained in a way that is practical, efficient, and sensitive to the needs of both users and services.

We hope that our work will increase the mainstream acceptance of anonymizing networks such as Tor, which has, thus far, been completely blocked by several services because of users who abuse their anonymity.

V. REFERENCES

[1] M., Bellare; H., Shi, and C., Zhang, (2005). "Foundations of Group Signatures: The Case of Dynamic Groups", Proceedings Cryptographer's Track at RSA Conference (CT-RSA), Springer, pp. 136-153.

[2] D., Boneh, and H., Shacham, (2004). "Group Signatures with Verifier-Local Revocation", Proceedings ACM Conference, Computer and Comm. Security, pp. 168-177.

[3] J., Camenisch, and A., Lysyanskaya, (2002). "Dynamic Accumulators and Application to Efficient Revocation of Anonymous Credentials", Proceedings Annual International Cryptology Conference (CRYPTO), Springer, pp. 61-76.

[4] J., Camenisch, and A., Lysyanskaya, (2004). "Signature Schemes and Anonymous Credentials from Bilinear Maps", Proceedings Annual International Cryptology Conference (CRYPTO), Springer, pp. 56-72.

[5] C., Cornelius; A., Kapadia; P., P. Tsang, and S., W. Smith, (2008). "Nymble: Blocking Misbehaving Users in Anonymizing Networks", Technical Report TR2008-637, Dartmouth College, Computer Science, December.

[6] J., Feigenbaum; A., Johnson, and P., F. Syverson, (2007). "A Model of Onion Routing with Provable Anonymity", Proceedings Conference Financial Cryptography, Springer, pp. 57-71.

[7] J., E. Holt, and K., E. Seamons, (2006). "Nym: Practical Pseudonymity for Anonymous Networks", Internet Security Research Lab Technical Report 2006-4, Brigham Young University, June.

[8] P., C. Johnson; A., Kapadia; P., P. Tsang, and S., W. Smith, (2007). "Nymble: Anonymous IP-Address Blocking", Proceedings Conference Privacy Enhancing Technologies, Springer, pp. 113-133.

[9] B., N. Levine; C., Shields, and N., B. Margolin, (2006). "A Survey of Solutions to the Sybil Attack", Technical Report 2006-052, University of Massachusetts, October.

[10] T., Nakanishi, and N., Funabiki, (2005). "Verifier-Local Revocation Group Signature Schemes with Backward Unlinkability from Bilinear Maps", Proceedings International Conference Theory and Application of Cryptology and Information Security (ASIACRYPT), Springer, pp. 533-548.

[11] 268 IEEE Transactions on Dependable and Secure Computing, Volume: 8, Number: 2, March-April, 2011.

*****

REDUCING PACKET LOSS IN INTER AND INTRA ASN HANDOVER IN MOBILE WIMAX NETWORK

Dhanam Mani [83]

ABSTRACT

The handoff is defined as the processes required for transferring the physical layer connectivity of a Mobile Node (MN) from one access point AP to another. In addition to physical connectivity it may also require transfer of some state information with respect to this MN. The existing GRAC provides a systematic way to solve the problem effectively and is fully compatible with the WiMAX standards, and can be used with other AC algorithms. We derive the performance bounds mathematically and show that the performance of the proposed GRAC approaches the lower bound.

The model analyzes the performance bounds of the system. Extensive simulations are conducted to validate the performance analysis. In this existing system during handoff the communication may be degraded or interrupted due to the packets loss. The main goal of this work was to optimize the network load and keeps packet loss at the required levels for VoIP communication. The work has been classified into two parts; the first part deals with inter ASN Handover. It uses GRAC Algorithm.

The second part deals with intra ASN handover .This part deals with investigation of the network conditions during the single path transmission, whereas the other investigates the wireless conditions during the multi-path transmission. Other two techniques are employed: Multi-Scan and Cross layer. Using the Multi-Scan approach, MN will be able to connect to two APs simultaneously using two interface cards. In addition, cross-layer approach enables the transport layer HM to obtain information about frame retransmission form the MAC layer.

Based on the MAC information about the number of frame retransmission, the handoff manager will decide either to send the packets via the single path transmission using single interface or the multi-path transmission using the two interfaces. This Combined inter and intra ASN Handover technique reduces the packet loss and improve the Mobile WiMAX system performance.

KEYWORDS

Mobile Node (MN), GRAC, VoIP, Algorithm, Multi-Scan and Cross layer, ASN etc.

I. INTRODUCTION

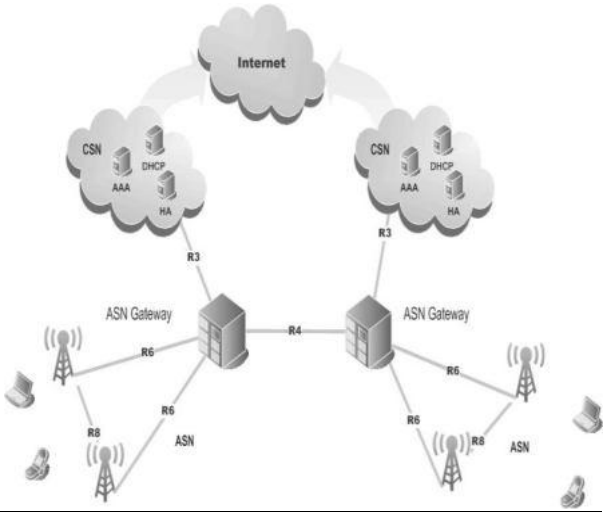
The IEEE 802.16 Working Group established by the IEEE Standards Board in 1999 has developed and published several versions of air interface standards for wireless metropolitan area networks (WMANs) with focus on medium access control (MAC) and PHY. While the initial versions of 802.16/a/d focused on fixed applications, the latest versions - 802.16-2005 (16e) amendment [1] and 802.16-REV2 - include many new features and functionalities needed to support enhanced quality of service (QoS) and mobility. The 802.16 Working Group is currently focusing on the specification for next-generation systems in the 802.16m Task Group. The WiMAX Forum was established in 2003 to promote and enable deployment of WiMAX as a new broadband access technology based on 802.16 standards.

WiMAX is a very useful technology in the suburban and rural blackout areas that currently have no broadband Internet access because phone and cable companies have not yet run the necessary wires to those remote locations. WiMAX forum developed the WiMAX Network Architecture. It is shown in the bellow fig 1. It has been launched in dozens of center around the nation. It provides benefits like highly improved bandwidth and speed over recent commercial cellular networks. A single WiMAX tower can provide coverage up to 3,000 square miles.

A WiMAX system consists of two parts:

- A **WiMAX tower**: It is similar in concept to a cell-phone tower - A single WiMAX tower can provide coverage to a very large area as big as 3,000 square miles (approximately 8,000 square km).
- A **WiMAX receiver**: The receiver and antenna could be a small box or PCMIA card, or they could be built into a laptop the way WiFi access is today.

Figure-1



Sources: Authors Compilation.

WiMAX Forum has defined two tired mobility management: **ASN Anchored Mobility and CSN Anchored Mobility:**

ASN Anchored Mobility refers to the procedures associated with the MS's movement between BSs, which may belong to the same or different ASN GWs. In ASN Anchored Mobility, the context of the designated MS is transferred from the previous BS to the new BS. Without performing CSN Anchored Mobility, ASN Anchored Mobility can minimize handover delay and packet loss.

[83] PG Scholar, Karpagam University, Tamil Nadu, India, dhanamani_raj@yahoo.com

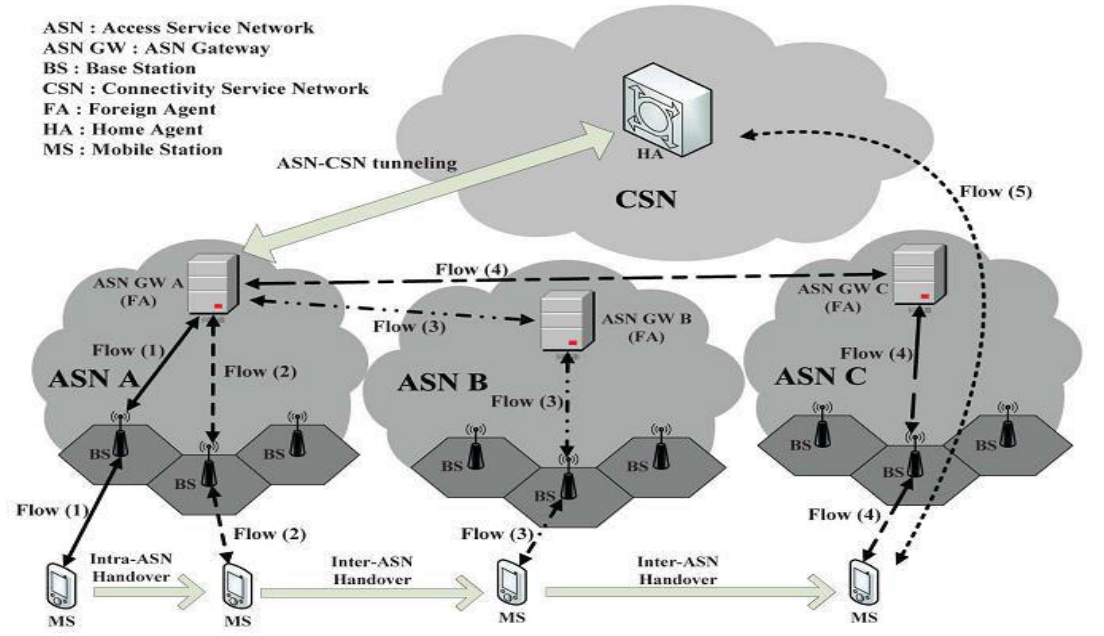
For example, an MS may perform intra-ASN handover (e.g., changing from Flow (1) to Flow (2) in Fig. 2) while still attaching to the same ASN GW. In addition, an MS may perform inter-ASN handover (e.g., changing from Flow (2) to Flow (3) in Fig. 2) where the ASN GW A is the traffic anchor point and responsible for ASN-CSN tunneling. That is, traffic is still sent to ASN GW A, which then further tunnels traffic to ASN GW B. In Flow (1) and Flow (2), the MS is called Serving MS of ASN GW A. In Flow (3), the MS is called Anchored MS of ASN GW A and handover MS of ASN GW B. In such case, the ASN GW A and ASN GW B are called anchored ASN GW and Serving ASN GW respectively.

CSN Anchored Mobility refers to the process of changing the traffic anchor point and is independent of the MS's link layer handover. It is also called ASN GW relocation. For example, if CSN Anchored Mobility is not performed, when the MS roams from ASN GW B to ASN GW C in Fig. 2, ASN GW A will tunnel traffic to ASN GW C. The MS is still served by two ASN GWs (ASN GW A and ASN GW C). As aforementioned discussion, the MS is called Anchored MS of ASN GW A.

Later on, the ASN GW A may request the MS to carry out CSN Anchored Mobility, i.e., ASN GW relocation. This may happen due to the heavy load of the ASN GW A, to reduce end-to-end latency, or for resource optimization purposes [4], [5]. After performing ASN GW relocation, the traffic anchor point is changed to ASN GW C. The MS then is not served by ASN GW A.

This is shown in Fig. 2 after changing from Flow (4) to Flow (5). Thus minimize handover delay and packet loss.

Figure-2



Sources: Authors Compilation.

II. EXISTING SYSTEM

In existing system goal is to design a stand-alone algorithm such that each ASN GW can determine when to request Anchored MSs to perform ASN GW relocation. The existing algorithm does not need to exchange information between neighboring ASN GWs. It also does not require centralized coordination and any assistance from extra servers. In addition, the proposed algorithm does not need to predict the movement of the mobile stations. It combines AC algorithm with a prediction technique to determine when is necessary to perform ASN GW relocation. Thus, it is called Gateway Relocation AC (GRAC). The GRAC consists of two components. The first one is AC algorithm. The prediction algorithm based on Wiener Process (WP) is then presented the ideas of AC algorithms are similar, although they have different names. Basically, the overall resources are partitioned and some resource is preserved for the handover MSs only. The proposed GRAC can work with any AC algorithm. In this section, they simply pick up the new call bounding algorithm. For simplicity, here they assumed that the resource assigned to each MS in one ASN GW is equal. The main point is not on a specific AC algorithm. The focus is on how to modify an AC algorithm for the two-tier mobility management in WiMAX.

Disadvantages

- The packet loss should be reduced.
- There is a need of effective algorithm to reduce the handoff mechanism.

III. PROPOSED SYSTEM

Handoff in WLANs has been addressed by many researchers, the main goal was to have seamless handoff by keeping delay, packet loss and jitter at acceptable levels to suit time sensitive applications such as VoIP. Layer 2 handoff schemes try to minimize the packet loss and delay on the MAC. Layer 3 handoff schemes try to come up with end-to-end handoff management schemes by changing the network infrastructure. The handoff phenomenon has been addressed throughout two main directions. One of them manages the aspects that are related to minimizing the packet loss and packet delay in both types of handoff, layer 2 and layer 3 handoffs, whereas the

second addresses the handoff criteria and the time when the handoff process should be triggered. Upper layer handoff managements rely on the end-to-end management schemes that introduce new hardware for managing the MN connectivity.

The handoff management scheme came up with a promising approach that uses the cross layer multi-scan along with the frame retransmission as a handoff trigging criterion. This approach eliminates the communication interruption, reduces packet loss, and keeps VoIP communication requirements at acceptable levels. However it introduces an extra network overhead during handoff **Network Load and Packet Loss Optimization during Handoff Using Multi-Scan Approach** due to the transmission of the same packet via both interfaces (multi-path transmission).

The work has been classified into two parts the first part deals with the investigation of the network conditions during the single path transmission, whereas the other investigates the wireless conditions during the multi-path transmission Figure2. During the single path transmission the HM keeps monitoring the link state by examining the number of frame retransmission in the MAC layer and compares it with a pre-defined Multi-Path Threshold (MPT).

The more closely the two values indicate a bad link condition and switching to multi-path has to occur. Instability threshold also introduced to control the faulty handoff indicators due to barriers and other wireless network interferences. During the multi-path transmission both links are investigated based on the number of frame retransmissions. Single Path Threshold (SPT) and stability counter are used as control parameters. In this case the link with the good condition will be used for next single path transmission phase.

Using Multi-scan approach intra network handover performance is improved. It is combined with inter network handover GRAC Algorithm. It improves the overall system performance.

Advantages

- Network load and packet loss can be reduced.
- Low construction cost.

IV. CONCLUSIONS

We proposed a GRAC algorithm which considers admission control and ASN GW relocation jointly to improve the performance of WiMAX networks during inter ANS Handover. The traditional AC algorithms cannot be used directly when the two-tiered mobility management is deployed in WiMAX because some MSs may be served by two ASN GWs. If there are many Anchored MSs, new incoming users will likely be rejected due to the lack of resources.

In the proposed GRAC, the AC algorithm cooperates with the ASN GW relocation. When a new MS arrives and there is no resource for the newly arrived MS, the proposed GRAC will request an Anchored MS to perform ASN GW relocation. In addition, for handover MSs, the WP-based prediction algorithm can trigger the ASN GW relocation at an appropriate time.

It can also estimate how many Anchored MSs should be relocated. Also in the intra ASN Handover is handled by using Multi scan Approach. The numerical results show that the proposed algorithms can effectively reduce blocking probability, dropping probability, and average signaling overhead. It also increases the average serving rate.

V. REFERENCES

[1] IEEE 802.16-2004 Std., Air Interface for Fixed Broadband Wireless Access Systems, IEEE, October, 2004.

[2] IEEE 802.16e-2005 Std., Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems-Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, IEEE, Feb. 2006.

[3] WiMAX Forum, http://www.wimaxforum.org, 2011.

[4] WiMAX Forum Std. 1.0, Rev. 4, WiMAX Forum Network Architecture (Stage 2: Architecture Tenets, Reference Model and Reference Points), WiMAX, Feb. 2009.

[5] WiMAX forum Std. 1.0, Rev. 4, WiMAX Forum Network Architecture (Stage 3: Detailed Protocols and Procedures), WiMAX, Feb. 2009.

[6] L., Nuaymi, WiMAX: Technology for Broadband Wireless Access.

[7] K., Etemad, (2008). "Overview of Mobile WiMAX Technology and Evolution", IEEE Comm. Mag., , October, Volume: 46, Number: 10, pp. 31-40.

[8] Y., Fang, and Y., Zhang, (2002). "Call Admission Control Schemes and Performance Analysis in Wireless Mobile Networks", IEEE Trans. Vehicular Technology, March, Volume: 51, Number: 2, pp. 371-382.

[9] D., Hong, and S., S. Rappaport, (1986). "Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Non Prioritized Handoff Procedures", IEEE Trans. Vehicular Technology, August, Volume: 35, Number: 3, pp. 77-92.

[10] A., C. Pang; Y., B. Lin; H., M. Tsai, and P., Agrawal, (2004). "Serving Radio Network Controller Relocation for UMTS All-IP Network", IEEE Journal Selected Areas in Comm., May, Volume: 22, Number: 4, pp. 617-629.

[11] J., Yao; J., W. Mark; T., C. Wong; Y., H. Chew; K., M. Lye, and K., C. Chua, (2004). "Virtual Partitioning Resource Allocation for Multiclass Traffic in Cellular Systems with QoS Constraints", IEEE Trans. Vehicular Technology, May, Volume: 53, Number: 3, pp. 847-864.

[12] J., H. Yeh; J., C. Chen, and P., Agrawal, (2009). "Fast Intra-Network and Cross-Layer Handover (FINCH) for WiMAX and Mobile Internet", IEEE Trans. Mobile Computing, April, Volume: 8, Number: 4, pp. 558-574.

[13] Z., H. Liu; S., Y. Pan, and J., C. Chen, (2008). "Access Service Network (ASN) Gateway Relocation Algorithms in WiMAX Networks", Proceeding IEEE International Conference Comm. (ICC '08), May, pp. 2674-2679.

*****

HYSTERSIS CONTROL OF HIGH EFFICIENCY DC-DC CONVERTER WITH TWO INPUT SOURCES

G. Mohanram[84] R. Nagarajan[85]

ABSTRACT

The aim of this study is to develop a high-efficiency converter with two input power sources for a distributed power generation mechanism. The proposed converter can boost the varied voltages of different power sources in the sense of hybrid power supply to a stable output dc voltage for the load demand. An auxiliary circuit in the proposed converter is employed for achieving turn-ON zero-voltage switching (ZVS) of all switches.

According to various situations, the operational states of the proposed converter can be divided into two states including a single power supply and a dual power supply. In the dual power-supply state, the input circuits connected in series together with the designed pulsewidth modulation can greatly reduce the conduction loss of switches. In addition, the effectiveness of designed circuit topology and the ZVS properties are verified by experimental results, and the goal of high-efficiency conversion can be obtained.

KEYWORDS

DC-DC Converter, Hybrid Power Supply, High Efficiency Power Conversion, Zero Voltage Switching (ZVS) etc.

INTRODUCTION

In the order to protect thenatural environment on the earth, the development of clean energy [1]-[3] without pollution has the major representative role in the last decade. By accompanying the permission of Kyoto Protocol, clean energies, such as fuel cell (FC), photovoltaic (PV), wind energy, etc., have been rapidly promoted. Due to the electric characteristics of clean energies, the generated power is critically affected by the climate or has slow transient responses, and the output voltage is easily influenced by load variations. [4]. Thus, a storage element is necessary to ensure proper operation of clean energies. Batter-ies or supercapacitors are usually taken as storage mechanisms for smoothing output power, start-up transition, and various load conditions [5], [6]. The corresponding installed capacity of clean energies can be further reduced to save the cost of system purchasing and power supply. For these reasons, hybrid power conversion systems (PCS) have become one of interesting re-search topics for engineers and scientists at present.

In the conduction mode (DCM) is utilized for achieving turn-ON ZVS of all the switches, and the huge reverse-recovery current of the output diode in the traditional boost converter can be removed via the utilization of an auxiliary inductor in the various conections in the mode of series connected with a diode. Consequently, the proposed dual-input converter can efficiently convert two power sources with different voltages to a stable dc-bus voltage. According to the power dispatch, this converter could be operated at two states including a single power-supply state and a dual power-supply state. This study is organized into four sections. Following the topology and operation of the proposed high-efficiency dual-input converter are presented in Section II.

Based on power electronics technique, the diversely developed power conditioners including dc-dc converters and dc-ac inverters are essential components for clean-energy applications. Generally, one power source needs a dc-dc converter either for rising the input voltage to a certain band or for regulating the input voltage to a constant dc-bus voltage [6]-[8]. However, conventional converter structures have the disadvantages of large size, complex topology, and expensive cost. In order to simplify circuit topology, improve system performance, and reduce manufacturing cost, multi-input converters have received more attentions in recent years [9]-[18].

Liu and Chen [9] proposed a general approach for developing multi-input converters. By analyzing the topologies of converters, the method for synthesizing multi-input converters was inspired by adding an extra pulsating voltage or a current source to a converter with an appropriate connection. Waiet al. [11], [12] presented multi-input converters with high stepup ratios, and the goal of high-efficiency conversion was obtained. However, these topologies are not economic for the non Electrical components. Tao et al. [13] and Matsuo et al. [14] utilized multiwinding-type transformers to accomplish the power conversion target of multi-input sources. Although these topologies were designed based on time-sharing concept, the complexity of driving circuits will be increased by the control techniques in [13] and [14].

Marchesoni and Vacca [15] investigated a newly designed converter with the series connected input circuits to achieve the goal of multiple input power sources. The installation cost of the converter with few components was certainly reduced. The feature of [15] is that the conduction losses of switches can be greatly reduced, especially in the dual power-supply state. Unfortunately, the hard-switching problem and the huge reverse-recovery current within the output diode degrade the conversion efficiency as a traditional boost converter [6].

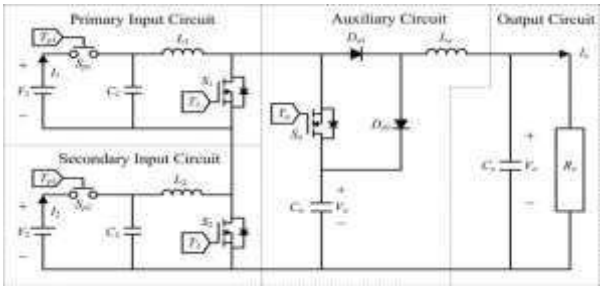
Kwasinski [16] discussed the evolution of multiple input converters from their respective single-input versions. Based on several assumptions, restrictions, and conditions, these analyses indicate some feasible and unfeasible frameworks for multiple input developments. Li et al. [17] investigated a set of basic rules for generating multiple-input converter topologies, and systematically generated two families of multiple-input converters.

Qian et al. [18] designed a novel converter topology with four power ports including two sources, one bidirectional storage port, and one isolated load port. The zero-voltage switching (ZVS) can be achieved for four main switches.

In this study, a high-efficiency ZVS dual-input converter is investigated, and this converter directly utilizes the current-source type applying to both the sources.

[84]PG Scholar, Raja Engineering college of Tehnology, Tamil Nadu, India, prabhu.klnce@gmail.com
[85]PG Scholar, Raja Engineering college of Tehnology, Tamil Nadu, India, bala_guys@yahoo.co.in

Figure-1: Circuit Topology of High-Efficiency Dual-Input Converter



Sources: Authors Compilation.

Based on the series-connected input circuits and the designed pulsewidth modulation (PWM) driving signals, the conduction loss of the switches can be greatly reduced in the dual power-supply state. Lee et al. [19] performed zero-current-transition dc-dc converters without additional current stress and conduction loss on the main switch during the resonance period of the auxiliary cell. The auxiliary cell provides zero-current-switching turn-OFF for all active switches and minimizes the reverse recovery problem of the main diode. The modified type of this representative auxiliary cell in [19] is introduced into the proposed dual-input converter to reduce the reverse-recovery currents of the diodes.

An auxiliary circuit with a small inductor operated in the discontinuous conduction mode (DCM) is utilized for achieving turn-ON ZVS of all the switches, and the huge reverse-recovery current of the output diode in the traditional boost converter can be removed via the utilization of an auxiliary inductor series connected with a diode.

Consequently, the proposed dual-input converter can efficiently convert two power sources with different voltages to a stable dc-bus voltage. According to the power dispatch, this converter could be operated at two states including a single power-supply state and a dual power-supply state. This study is organized into four sections.

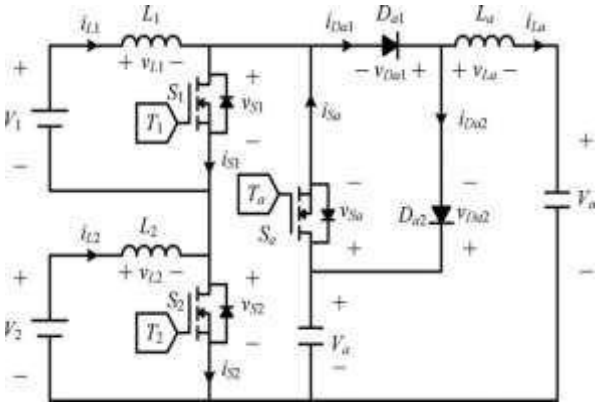
Following the introduction in Section I, the topology and operation of the proposed high-efficiency dual-input converter are presented in Section II. In Section III, experimental results are provided to validate the effectiveness of the proposed converter. Conclusions are drawn in Section IV.

II. TOPOLOGY AND OPERATION OF DUAL-INPUT CONVERTER

Fig. 1 shows the circuit topology of the proposed ZVS dual input converter. It contains four parts including a primary input circuit, a secondary input circuit, an auxiliary circuit, and an output circuit. The major symbol representations are summarized as follows. V_1 and I_1 denote the primary input voltage and current, respectively. V_2 and I_2 exhibit the secondary input voltage and current, respectively. S_{P1}, S_{P2}, T_{P1}, and T_{P2} express the power ON/OFF switches and their driving signals produced by the power management. C_i, L_i, S_i, and T_i (i = 1, 2) represent individual capacitors, inductors, switches, and driving signals in the primary and secondary input circuit, respectively. C_a, L_a, D_{a1}, and D_{a2} are the auxiliary capacitor, inductor, and diodes of the auxiliary circuit. S_a and T_a are the auxiliary switch and its driving signal, which is generated by the PWM. C_o, V_o, I_o, and R_o describe the output capacitor, voltage, current, and equivalent load, respectively.

For the convenience of analyses, the simplified equivalent circuit is depicted in Fig. 2, and the directional definition of significant voltages and currents are labeled in this figure. The simplification in Figure-2 is compliant with the following assumptions: 1) all power switches and diodes have ideal characteristics without considering voltage drops when these devices are conducted; 2) the capacitors C_a and C_o are large enough so that the voltage ripples due to switching are negligible and could be taken as constant voltage sources V_a and V_o; and 3) the power ON/OFF switches S_{P1} and S_{P2} are omitted.

Figure-2: Equivalent Circuit



Sources: Authors Compilation.

According to different power conditions, the operational states of the proposed converter can be divided into two states including a single power-supply state with only one input power source and a dual power-supply state with two input power sources. The powers produced by the voltage sources V_1 and V_2 are referred as P_1 and P_2 , respectively, while the power consumed by the load is referred as P_o .

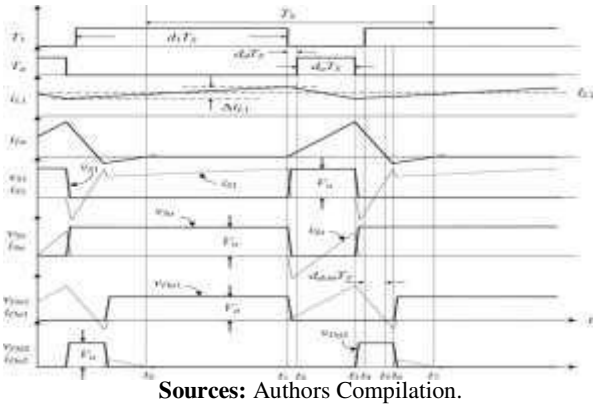
If the condition of $P_1 > P_o$ ($P_2 > P_o$) holds, the switch S_{P1} (S_{P2}) turns ON to supply the power with a single input power source V_1 (V_2). On the contrary, the switches S_{P1} and S_{P2} turn ON to supply the power with two input power sources if the conditions of $P_1 > P_o$ and $P_2 > P_o$ fail. The detailed operational stages are described as follows:

A. Single Power-Supply State

By turning off one power ON/OFF switch S_{P1} or S_{P2} for cutting off the connection between the power source and the converter, the other input power source V_2 or V_1 can supply alone for supporting the output demand. The primary input power supply is considered, for example, to explain how to operate in this state, i.e., the switch S_{P2} is always turned OFF and the switch S_2 is triggered all the while for minimizing the conduction loss.

The switching period is defined as T_S. d_1 and d_a denote the duty cycles of the switch S_1 and S_a, respectively. d_d and d_{dcm} present the duty cycles of the dead time and the freewheeling time of the auxiliary inductor. Note that the auxiliary inductor is designed to operate in the DCM. The characteristic wave forms and topological modes of the single power-supply state are depicted in Figs. 3 and 4, respectively. The complete operation modes in a switching period of the converter are discussed as follows:

Figure-3: Characteristic Wave Forms in Single Power-Supply State



Sources: Authors Compilation.

Mode 1 [t_0 -t_1]: At t_0, the auxiliary inductor current $i_{L\,a}$ returned to zero. The switch S_1 is continuously conducted and the auxiliary switch S_a is still turned OFF. The primary inductor L_1 is linearly charged by the primary input voltage V_1. The auxiliary switch voltage $v_{S\,a}$ is equal to the auxiliary capacitor voltage V_a.

Mode 1 [t_0 -t_1]: At t_0, the auxiliary inductor current $i_{L\,a}$ returned to zero. The switch S_1 is continuously conducted and the auxiliary switch S_a is still turned OFF. The primary inductor L_1 is linearly charged by the primary input voltage V_1. The auxiliary switch voltage $v_{S\,a}$ is equal to the auxiliary capacitor voltage V_a.

Mode 2 [t_1 -t_2]: At t_1 , the switch S_1 is turned OFF, the switch voltage $v_{S\,1}$ is rising to the auxiliary capacitor voltage V_a , and the auxiliary switch voltage $v_{S\,a}$ is decreasing to zero. The body diode of the auxiliary switch S_a is conducted for receiving the primary inductor current $i_{L\,1}$ to charge the auxiliary capacitor. Therefore, switch current i_{Sa} is negative. Besides, the auxiliary inductor current linearly increases, and its slope is dependent on the auxiliary inductor voltage $v_{L\,a}$, which is equal to V_a–V_o. Continuously, the primary auxiliary diode $D_{a\,1}$ is conducted.

Mode 3 [t_2 -t_3]: At t_2, the auxiliary switch S_a is turned ON with ZVS because the body diode has been already conducted for carrying the primary inductor current $i_{L\,1}$. After the auxiliary inductor current $i_{L\,a}$ increases to be larger than the primary inductor current $i_{L\,1}$, the auxiliary switch current $i_{S\,a}$ becomes positive. The discharging current from the auxiliary capacitor together with the primary inductor current $i_{L\,1}$ releases the stored energy to the output voltage V_o. During modes 2 to 3 ($t = t_1 - t_3$), the time interval can be written as $(d_d + d_a)\,T_S$. The auxiliary inductor current $i_{L\,a}$ and the primary inductor current:

$$i_{L\,a}(t) = \frac{(V_a - V_o)(t - t^1)}{L_a} \qquad (1)$$

$$i_{L\,1}(t) = (I_{L\,1} + 0.5\Delta i_{L\,1}) + \frac{(V_1 - V_a)(t - t^1}{L_1} \qquad (2)$$

Where $I_{L\,1}$ is the average value of the primary inductor current i_{L1}, and Δi_{L1} is the corresponding peak-to-peak current ripple. Note that, the time interval (t_1 -t_2) in mode 2 is extremely short so that it could be regarded as the same time in Fig. 4. At t_3, the maximum values of auxiliary inductor current i_{La} can be calculated as:

$$i_{L\,a}(t_3) = \frac{(V_a - V_o)(d_d + d_a)T_S}{L_a} \qquad (3)$$

According to (2), the current ripple $\Delta i_{L\,1}$ can be rewritten as

$$\Delta i_{L\,1} = \frac{(V_a - V_1)(d_d + d_a)T_S}{L_1} \qquad (4)$$

Mode 4 [t_3 -t_4]: At t_3 , the auxiliary switch S_a is turned OFF, because the auxiliary inductor current $i_{L\,a}$ is greater than the primary inductor current i_{L1}, the parasitic capacitor of the auxiliary switch S_a is charged by the auxiliary inductor current $i_{L\,a}$ so that the auxiliary switch voltage $v_{S\,a}$ rises. At the same time, the energy stored in the parasitic capacitor of the switch S_1 will release to the output voltage V_o via the inductor current i_{La} so that the switch voltage $v_{S\,1}$ decreases. The switch current i_{Sa} falls down to zero and the switch voltage $v_{S\,a}$ rises to the auxiliary capacitor voltage V_a . The body diode of the switch S_1 is conducted for carrying the differential current without strain. Besides, the auxiliary inductor voltage v_{La} is equal to $-V_o$, and the current i_{La} linearly decreases. The energy stored in the auxiliary inductor L_a starts to discharge into the output voltage V_o as freewheeling.

Mode 5 [t_4 -t_5]: At t_4, the switch S_1 is turned ON with ZVS upon the condition that the auxiliary inductor current i_{La} is still larger than the primary inductor current i_{L1}. The auxiliary inductor current i_{La} continuously decreases with the slope $-V_o/L_a$. After the current i_{La} is smaller than the primary inductor current i_{L1}, the switch current i_{S1} is positive. During modes 4 to 5 ($t = t_3 - t_5$), the time interval can be written as $(d_d + d_{dcm})T_S$. The auxiliary inductor current i_{La} and the primary inductor current $i_{L\,1}$ can be expressed as:

$$i_{L\,a}(t) = \frac{[(V_a - V_o)(d_d + d_a)T_S - V_o(t - t^3)]}{L_a} \qquad (5)$$

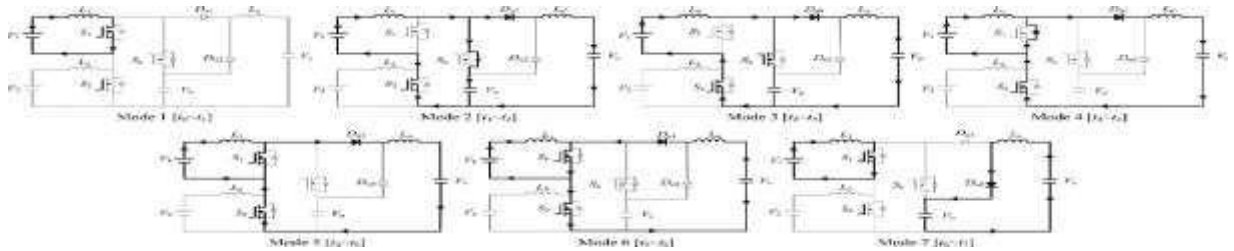
$$i_{L\,1}(t) = (I_{L\,1} - 0.5\Delta i_{L\,1}) + \frac{V^1(t - t^3)}{L_1} \qquad (6)$$

Mode 6 [t_5-t_6]: At t_5, the auxiliary inductor current i_{La} is equal to zero. Substituting $i_{La}(t_5) = 0$ into (7), the relation between the voltages V_a and V_o can be derived as:

$$(V_a - V_o)(d_d + d_a) = V_o(d_d + d_{dcm}). \qquad (7)$$

In this mode, the parasitic capacitor of the primary auxiliary diode D_{a1} is charged by the output voltage V_o with a small reverse-recovery current.

Figure-4: Topological Modes In Single Power-Supply State



Sources: Authors Compilation.

Mode 7 [t_6 -t_7]: At t_6, the diode voltage v_{Da1} is rising to the output voltage V_o, the secondary auxiliary diode D_{a2} is conducted for receiving the auxiliary inductor current i_{La} to charge the auxiliary capacitor voltage V_a, and then the auxiliary inductor current i_{La} returns to zero. In the single power-supply state with the primary input power, the switch S_{P2} is always turned OFF and the switch S_2 is triggered all the while. It means that the switch S_2 works as a synchronous rectifier for avoiding the current to flow through its body diode and reducing the power losses in modes 2-7 in Figure-4.

According to volt-second balance theory [20], the voltage-second production of the primary inductor L_1 in a switching period should be equal to zero. Thus, one can obtain:

$$V_1 (d_1 + d_d)T_S + (V_1 - V_a)(d_a + d_d)T_S = 0 \quad (8a)$$

$$V_1 = V_a (d_a + d_d). \quad (8b)$$

Assume that the dead-time duty cycle d_d is much smaller than the duty cycle of the switch d_1 ; the summation of the duty cycles d_1 and d_a approaches to 1. The relationships of (7) and (8b) can be represented as:

$$V_o = \frac{(1 - d_1)V_a}{(1 + d_{dcm} - d_1)} \quad (9a)$$

$$V_1 = (1 - d_1)V_a \quad (9b)$$

$$\frac{V_o}{V_1} = \frac{1}{(1 + d_{dcm} - d_1)}. \quad (9c)$$

Because the average current of the output capacitor C_o should be zero over a switching period for a constant output voltage V_o, the balance equation can be expressed via (3) as:

$$\frac{0.5(V_a - V_o)(1 - d_1)T_S(1 - d_1 + d_{dcm})}{L_a} = \frac{V_o}{R_o}. \quad (10)$$

From the algebraic operation via (9) and (10), the duty cycle and the voltage gain of the converter can be derived as:

$$d_{dcm} = 0.5(1 - d_1)\left[\sqrt{1 + \frac{8L_a}{R_o T_S (1 - d_1)^2}} - 1\right] \quad (11a)$$

$$V_o = \frac{2V_1}{(1 - d_1)\left[1 + \sqrt{1 + \frac{8L_a}{R_o T_S (1 - d_1)^2}}\right]}. \quad (11b)$$

By the similar derivation process, the voltage gain of the single power-supply state with the secondary input power source also can be represented as:

$$V_a = \frac{V_2}{(1 - d_2)} \quad (12a)$$

$$V_o = \frac{2V_2}{(1 - d_2)\left[1 + \sqrt{1 + \frac{8L_a}{R_o T_S (1 - d_2)^2}}\right]} \quad (12b)$$

Where d_2 denotes the duty cycle of the switch S_2 .

B. Dual Power-suply State

When the proposed converter is operated in the dual power-supply state with two input power sources, it can be taken as a superposition process of the primary and secondary input circuits. In this state, the summation of duty cycles d_1 and d_2 should be greater than 1, i.e., each of duty cycles d_1 and d_2 is:

(9a): Securely greater than 0.5. Moreover, the symbols d_{a1} and d_{a2} denote the first and the second duty cycles of the switch S_a

(9b): d_{dcm1} and $d_{dcm\,2}$ present the first and the second duty cycles of the freewheeling times of the auxiliary inductor

(9c): The auxiliary inductor is also designed to operate in the DCM

In order to explain the operational principle in the dual power-supply state easily, the following theoretical analysis is based on the assumption of $i_{L\,1} > i_{L\,2} > |i_{L\,1} - i_{L\,2}|$, where $|\cdot|$ is the absolute operator. The characteristic waveforms and topological modes of the dual power-supply state are depicted in Figs. 5 and 6, respectively. Note that the time intervals in modes 2, 4, 9, and 11 are extremely short so that each interval could be regarded as the same time in Fig. 5. The operation modes in this state are discussed as follows.

Mode 1 [t_0 -t_1]: At t_0, the auxiliary inductor current i_{La} returned to zero. The switches S_1 and S_2 are continuously conducted. The auxiliary switch S_a is still turned OFF. They appear complementarily. By observing the switch voltage v_{S1} and current i_{S1}, the characteristic of turning on with ZVS is obvious due to the current is negative before the switch is turned ON. Fig. 7(c) shows the driving signals (T_1 and T_a), the switch voltage v_{Sa}, and current i_{Sa}. The ZVS turn ON of the switch S_a can also be achieved. The auxiliary capacitor C_a receiving the primary inductor current i_{L1} at first, then the energy is released to the output terminal by the positive current $i_{S\,a}$. Fig. 7(d) illustrates the driving signals (T_1 and T_a), the prime auxiliary diode voltage $v_{D\,a\,1}$, and current $i_{D\,a\,1}$. It states that the diode current i_{Da1} climbs with the slope $(V_a - V_o)/L_a$ and falls with the slope V_o/L_a.

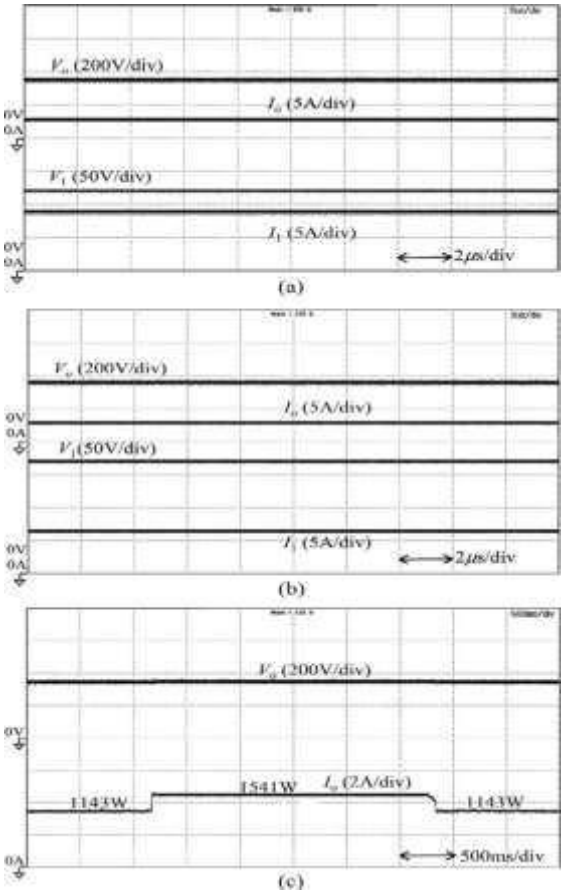
Moreover, the phenomenon of a huge reverse recovery current disappears via the utilization of an auxiliary inductor series onnected with a diode when the diode current falls to zero in comparison with the traditional step-up converters. The driving signals (T_1 and T_a), the second auxiliary diode voltage v_{Da2}, and current i_{Da2} are depicted in Fig. 7(e). The second auxiliary diode D_{a2} is conducted for receiving the auxiliary inductor current i_{La} to charge the auxiliary capacitor. Fig. 7(f) shows the input/output voltage and current waveforms. When the input voltage is V_1 = 170 V, the output voltage V_o can be controlled to 360 V and the maximum output voltage ripple is less than 3.5 V (0.97%).

Fig. 8 shows the experimental input/output voltage and current responses of the dual-input converter at the single power-supply state due to different input voltages and varied output powers, where the response of V_1 = 120 V with 1143 W output power is depicted in Fig. 8(a); the response of V_1 = 170 V with 1143 W output power is depicted in Fig. 8(b); the response of the output power variation between 1143 and 1541 W is depicted in Fig. 8(c), From the measurable data, the conversion efficiency due to the load variation between 1143 and 1541 W is varied between 97% and 95.8% in Fig. 8(c). As can be seen from this

figure, the robustness of the proposed dual-input converter with the closed-loop PI voltage control under different input voltages and varied output powers is obvious. Fig. 9 performs the conversion efficiency at the single power-supply state with the closed-loop PI voltage control. The measured conditions are set with a primary input power source of V_1 = 170 V and a constant output voltage of V_o = 360 V. From the experimental results, the maximum efficiency is measured to be about 97% due to the ZVS property of all switches.

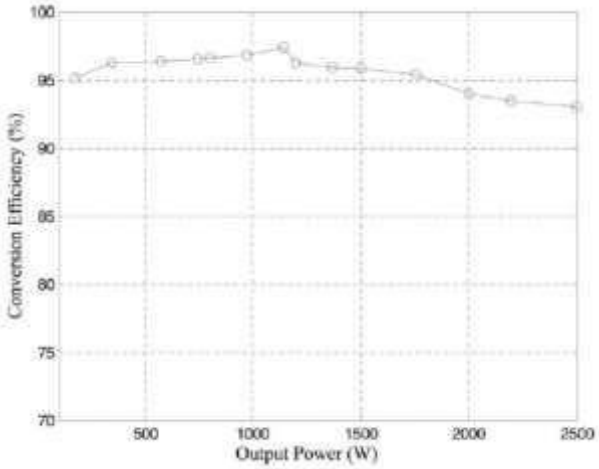
By considering the possible applications of the proposed dual input converter with the constant current control, Fig. 10 shows the experimental results at the single power-supply state with the closed-loop PI current control. In this experiment, the test conditions are set with V_2 = 170 V, 2.1-kW output power, and the desired input current I_2 = 13 A. In practical applications, the current of the input power source (not the inductor current) is sensed for the control utilization because the source current is smoother than the inductor current with charge/discharge slope. The proposed converter under the closed-loop current control indeed produces a constant input current of I_2 = 13 A via the DSP module written within a PI feedback control law. Fig. 10(a) presents the driving signal T_2, the auxiliary capacitor voltage V_a, the output voltage V_o, and the secondary inductor current i_{L2}. The auxiliary capacitor voltage V_a is nearly a constant.

Figure-8: (a)-(c) Experimental input/output voltage and current responses of dual-input converter at single power-supply state with closed-loop voltage control due to varied input voltage and output power. (a) V_1 = 120 V with 1143-W output power, (b) V_1 = 170 V with 1143-W output power, and (c) output power variation between 1143 and 1541 W.



Sources: Authors Compilation.

Figure-9: Conversion efficiency at single power-supply state with closed-loop voltage control and V_1 = 170 V.



Sources: Authors Compilation.

CONCLUSIONS

This study has successfully developed a ZVS dual-input converter with hybrid power sources. The effectiveness of this converter is also verified by the experimental results. In the single power-supply state, the property of ZVS turn ON of all switches guarantees that switching losses can be reduced. In the dual power-supply state, the conduction loss can be effectively reduced by topological design of series connection of two input circuits. Besides, the reverse-recovery currents of the diodes are slight as well as the switching losses of the switches are effectively reduced. The maximum efficiency of the proposed converter operated in both operational states is higher than 95%. This new converter topology provides designers with an alternative choice to simultaneously convert hybrid power sources. In addition, the proposed high-efficiency dual-input converter also can work well in high-power level applications because the switching losses can be greatly reduced due to the ZVS property.

In general, the ground leakage current would be harmful to high-power nonisolated PV applications due to the presence of a parasitic capacitance between the PV cells and the metal frame of the PV panel, usually connected to earth. Xiao and Xie [21] focused on the leakage current suppressing method by considering all common-mode paths and presented a new full-bridge-type converter structure and a compensation strategy for half-bridge-type inverter. Yu *et al.* [22] studied a two-stage PV ac-module application including a nonisolated high step-up dc/dc converter and a newly designed H6-type dc/ac inverter to feature high efficiency over a wide load range, low ground leakage current, no need for split capacitors, and low-output ac-current distortion. If the proposed ZVS dual-input converter in this study is used for nonisolated PV applications, the ground leakage current issue due to the high-frequency voltage swing between the two negative terminals of the input sources also may cause safety and electromagnetic interference (EMI) problems because the input power sources do not share the same ground. These problems could be solved by the adoption of an EMI filter [21] for a dc power supply application or the integration with advanced inverter topologies [22] for an ac-module application in the future research.

REFERENCES

[1] S., Al-Hallaj, (2004). "More than Enviro-Friendly: Renewable Energy is also good for the Bottom Line", IEEE Power Energy Magazine, May/June, Volume: 2, Number: 3, pp. 16-22.

[2] P., Fairley, (2005). "The Greening of GE", IEEE Spectrum, July, Volume: 42, Number: 7, pp. 28-33.

[3] R., C. Dugan; T., S. Key, and G., J. Ball, (2006). "Distributed Resources Standards", IEEE Ind. Appl. Mag., January / February, Volume: 12, Number: 1, pp. 27-34.

[4] B., Yang; W., Li; Y., Zhao, and X., He, (2010). "Design and Analysis of a grid-Connected Photovoltaic Power System", IEEE Trans. Power Electron, April, Volume: 25, Number: 4, pp. 992-1000.

[5] S., K. Kim; J., H. Jeon; C., H. Cho; J., B. Ahn, and S., H. Kwon, (2008). "Dynamic Modeling and Control of a Grid Connected Hybrid Generation System with Versatile Power Transfer", IEEE Trans. Ind. Electron., April, Volume: 55, Number: 4, pp. 1677-1688.

[6] M., B. Camara; H., Gualous; F., Gustin, and A., Berthon, (2008). "Design and New Control of DC/DC Converter to Share Energy between Supercapacitors and Batteries in Hybrid Vehicles", IEEE Trans. Veh. Technol., vol. 57, no. 5, pp. 2721-2735, Sep. 2008.

[7] T., Bhattacharya; V., S. Giri; K., Mathew, and L., Umanand, (2009). "Multiphase Bi-Directional Flyback Converter Topology for Hybrid Electric Vehicles", IEEE Trans. Ind. Electron., January, Volume: 56, Number: 1, pp. 78-84.

[8] S., H. Park; S., R. Park; J., S. Yu; Y., C. Jung, and C., Y. Won, (2010). "Analysis and Design of a Soft-Switching Boost Converter with an HI-bridge Auxiliary Resonant Circuit", IEEE Trans. Power Electron., August, Volume: 25, Number: 8, pp. 2142-2149.

[9] Y., C. Liu; and Y., M. Chen, (2009). "A systematic Approach to Synthesizing Multi-Input DC-DC Converters", IEEE Trans. Power Electron., January, Volume; 24, Number: 1, pp. 116-127.

[10] Y., M. Chen; Y., C. Liu; S., C. Hung, and C., S. Cheng, (2007). "Multi-Input Inverter for Grid-Connected Hybrid PV/Wind Power System", IEEE Trans. Power Electron, May, Volume: 22, Number: 3, pp. 1070-1077.

[11] R., J. Wai; C., Y. Lin; L., W. Liu,, and Y., R. Chang, (2007). "High-Efficiency Single-Stage Bidirectional Converter with Multi-Input Power Sources", Inst. Electr. Eng. Proc.: Electr. Power Appl., September, Volume: 1, Number: 5, pp. 763-777.

[12] R., J. Wai; C., Y. Lin, and Y., R. Chang, (2009). "High Step-up Bidirectional Isolated Converter with Two Input Power Sources", IEEE Trans. Ind. Electron, July, Volume: 56, Number: 7, pp. 2629-2643.

[13] H., Tao; J., L. Duarte, and M., A. M. Hendrix, (2008). "Three-port triple-half-Bridge bidirectional Converter with zero-Voltage Switching", IEEE Trans. Power Electron., March, Volume: 23, Number: 2, pp. 782-792.

[14] H., Matsuo; W., Z. Lin; F., Kurokawa; T., Shigemizu, and N., Watanabe, "Characteristics of the multiple-input DC-DC converter", IEEE Trans. Ind. Electron., June, Volume: 51, Number: 3, pp. 625-631.

[15] M., Marchesoni, and C., Vacca, (2007). "New DC-DC Converter for Energy Storage System Interfacing in Fuel Cell Hybrid Electric Vehicles", IEEE Trans. Power Electron., January, Volume: 22, Number: 1, pp. 301-308.

[16] A., Kwasinski, (2009). "Identification of Feasible Topologies for Multiple-Input DC-DC Converters", IEEE Trans. Power Electron, March, Volume: 24, Number: 3, pp. 856-861.

[17] Y., Li. X. Ruan; D., Yang; F., Liu, and C., K. Tse, (2010). "Synthesis of Multiple-Input DC/DC Converters", IEEE Trans. Power Electron, September, Volume: 25, Number: 9, pp. 2372-2385.

[18] Z., Qian; O., Abdel-Rahman, and I., Batarseh, (2010). "An Integrated Four-Port DC/DC Converter for Renewable Energy Applications", IEEE Trans. Power Electron, July, Volume: 25, Number: 7, pp. 1877-1887.

[19] D., Y. Lee; M., K. Lee; D., S. Hyun, and I., Choy, (2003). "New zero-current-transition PWM DC/DC Converters without Current stress", IEEE Trans. Power Electron., January, Volume: 18, Number: 1, pp. 95-104.

[20] N., Mohan; T., M. Undeland, and W., P. Robbins, (1995). Power Electronics: Converters, Applications, and Design, New York: Wiley.

[21] H., Xiao, and S., Xie, (2010). "Leakage Current Analytical Model and Application in Single-Phase Transformerless Photovoltaic Grid-Connected Inverter", IEEE Trans. Electromagn. Compat., November, Volume: 52, Number: 4, pp. 902-913.

[22] W., Yu; J., S. Lai; H., Qian, and C. Hutchens, (2011). "High-efficiency MOSFET Inverter with H6-type Configuration for Photo Voltaic Nonisolated AC-module Applications", IEEE Trans. Power Electron., April, Volume: 26, Number; 4, pp. 1253-1260.

Note -Figure 5 to 7 is omitted from the paper.

*****

SIMULATION AND ANALYSIS OF NEW MULTI LEVEL INVERTER TOPOLOGY

S. Saipadhma[86]

ABSTRACT

Multilevel inverters have been widely accepted for high-power high-voltage applications. Their performance is highly superior to that of conventional two-level inverters due to reduced harmonic distortion, lower electromagnetic interference, and higher dc link voltages.

However, it has some disadvantages such as increased number of components, complex pulse width modulation control method, and voltage-balancing problem. In this paper, a new topology with a reversing-voltage component is proposed to improve the multilevel performance by compensating the disadvantages mentioned. This topology requires fewer components compared to existing inverters (particularly in higher levels) and requires fewer carrier signals and gate drives.

Therefore, the overall cost and complexity are greatly reduced particularly for higher output voltage levels. Finally, a prototype of the seven-level proposed topology is built and tested to show the performance of the inverter by experimental results.

KEYWORDS

Multilevel Inverter, Signal, Power Electronics, Voltage, SPWM, Topology etc.

INTRODUCTION

Multilevel power conversion was first introduced more than two decades ago. The general concept involves utilizing a higher number of active semiconductor switches to perform the power conversion in small voltage steps. There are several advantages to this approach when compared with the conventional power conversion approach. The smaller voltage steps lead to the production of higher power quality waveforms and also reduce voltage (dv/dt) stress on the load and the electromagnetic compatibility concerns [1]. Another important feature of multilevel converters is that the semiconductors are wired in a series-type connection, which allows operation at higher voltages. However, the series connection is typically made with clamping diodes, which eliminates overvoltage concerns. Furthermore, since the switches are not truly series connected, their switching can be staggered, which reduces the switching frequency and thus the switching losses. One clear disadvantage of multilevel power conversion is the higher number of semiconductor switches required. It should be pointed out that lower voltage rated switches can be used in the multilevel converter and, therefore, the active semiconductor cost is not appreciably increased when compared with the two level cases.

However, each active semiconductor added requires associated gate drive circuits and adds further complexity to the converter mechanical layout. Another disadvantage of multilevel power converters is that the small voltage steps are typically produced by isolated voltage sources or a bank of series capacitors. Isolated voltage sources may not always be readily available, and series capacitors require voltage balancing [2]. To some extent, the voltage balancing can be addressed by using redundant switching states, which exist due to the high number of semiconductor devices. However, for a complete solution to the voltage-balancing problem, another multilevel converter may be required [3]. In recent years, there has been a substantial increase in interest to multilevel power conversion. Recent research has involved the introduction of novel converter topologies and unique modulation strategies.

However, the most recently used inverter topologies, which are mainly addressed as applicable multilevel inverters, are cascade converter, neutral-point clamped (NPC) inverter, and flying capacitor inverter. There are also some combinations of the mentioned topologies as series combination of a two-level converter with a three-level NPC converter which is named cascade 3/2 multilevel inverter [4]. There is also a series combination of a three-level cascade converter with a five-level NPC converter which is named cascade 5/3 multilevel inverter [5]. Some applications for these new converters include industrial drives [6], flexible ac transmission systems (FACTS) [7]–[9], and vehicle propulsion [10], [11]. One area where multilevel converters are particularly suitable is that of renewable photovoltaic energy that efficiency and power quality are of great concerns for the researchers [12].

Some new approaches have been recently suggested such as the topology utilizing low-switching-frequency high-power devices [13]. Although the topology has some modification to reduce output voltage distortion, the general disadvantage of this method is that it has significant low-order current harmonics. It is also unable to exactly manipulate the magnitude of output voltage due to an adopted pulse width modulation (PWM) method [14]. In [15] and [16], the multilevel output is generated with a multi winding transformer.

However, the design and manufacturing of a multi winding transformer are difficult and costly for high-power applications. A novel four-level inverter topology is also proposed, and it is valid for inverters with even number of voltage levels and not capable of outputting a zero-voltage state. As a result, the inverter output phase voltage for zero modulation indexes is a bipolar waveform taking two distinct values and exhibits high rms value and considerable harmonic energy concentrated at the switching frequency. This is a disadvantage of the proposed inverter, particularly when it should output low or zero voltage to a load [17]. Another approach is selection based on a set target which can be either the minimum switches used or the minimum used dc voltage. It also requires different voltage source values which are defined according to the target selection [18].

However, this approach also needs basic units which are connected in series, and the basic units still require more switches than the proposed topology. Another disadvantage of the topology is that the power switches and diodes also need to have a different rating which is a major drawback of the topology. In [19], the voltage sources are not used efficiently in generating output voltage levels. For example, the topology in

[86]Student, MNM Jain Engineering College, Tamil Nadu, India, sai631981@yahoo.co.in

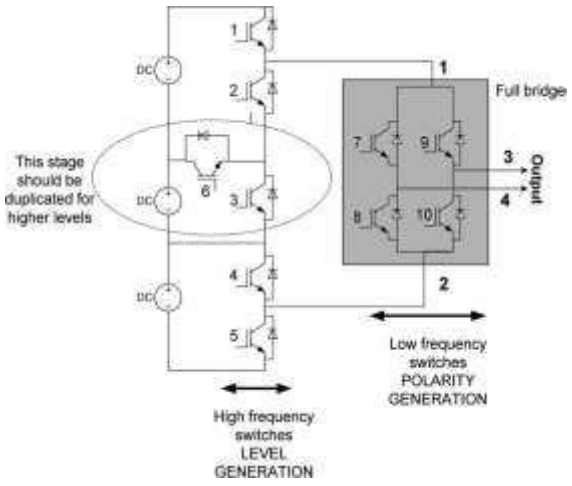
[19] can generate only five output levels with four dc sources, while conventional multilevel inverters can generate up to nine levels with the same number of power supplies. The proposed topology is a symmetrical topology since all the values of all voltage sources are equal. However, there are asymmetrical topologies [20] which require different voltage sources. This criterion needs to arrange dc power supplies according to a specific relation between the supplies.

Difference in ratings of the switches in the topology is also a major drawback of the topology. This problem also happens in similar topologies [21]–[23], while some of the high-frequency switches should approximately withstand the maximum overall voltage which makes its application limited for high-voltage products. In [24], a new approach has been proposed that decreases the number of required dc supplies and inserting transformer instead.

The main disadvantage of the approach is adding so many transformer windings which will add up to the overall volume and cost of the inverter. There is also another topology which requires more switches than the proposed topology for the same number of levels [25]. Some of the proposed topologies suffer from complexities of capacitor balancing [26]–[28]. In [26], the capacitor values used in the topology are proportional to the load current, and as the load current increases, a larger capacitor should be selected. In [27], the capacitor voltage will affect the output voltage when modulation index reaches near its extreme values, i.e., zero or one.

This paper presents an overview of a new multilevel inverter topology named reversing voltage (RV). This topology requires less number of components compared to conventional topologies. It is also more efficient since the inverter has a component which operates the switching power devices at line frequency. Therefore, there is no need for all switches to work in high frequency which leads to simpler and more reliable control of the inverter. This paper describes the general multilevel inverter schematic. A general method of multilevel modulation phase disposition (PD) SPWM is utilized to drive the inverter and can be extended to any number of voltage levels. The simulation and experimental results of the proposed topology are also presented.

Figure-1: Schematic of Seven-Level Inverter in Single Phase



Sources: Authors Compilation.

NEW MULTILEVEL TOPOLOGY

A. General Description

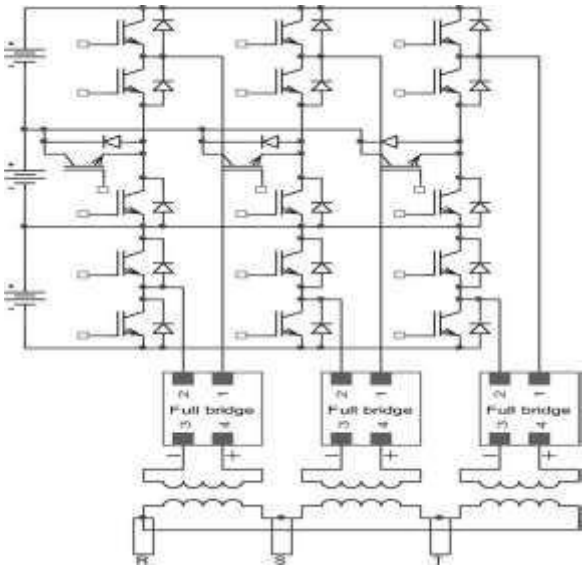
In conventional multilevel inverters, the power semiconductor switches are combined to produce a high-frequency waveform in positive and negative polarities. However, there is no need to utilize all the switches for generating bipolar levels. This idea has been put into practice by the new topology. This topology is a hybrid multilevel topology which separates the output voltage into two parts:

One part is named *level generation* part and is responsible for level generating in positive polarity. This part requires high-frequency switches to generate the required levels. The switches in this part should have high-switching-frequency capability. The other part is called *polarity generation* part and is responsible for generating the polarity of the output voltage, which is the low-frequency part operating at line frequency.

The topology combines the two parts (high frequency and low frequency) to generate the multilevel voltage output. In order to generate a complete multilevel output, the positive levels are generated by the high-frequency part (level generation), and then, this part is fed to a full-bridge inverter (polarity generation), which will generate the required polarity for the output. This will eliminate many of the semiconductor switches which were responsible to generate the output voltage levels in positive and negative polarities. The RV topology in seven levels is shown in Fig. 1. As can be seen, it requires ten switches and three isolated sources.

The principal idea of this topology as a multilevel inverter is that the left stage in Fig. 1 generates the required output levels (without polarity) and the right circuit (full-bridge converter) decides about the polarity of the output voltage. This part, which is named polarity generation, transfers the required output level to the output with the same direction or opposite direction according to the required output polarity. It reverses the voltage direction when the voltage polarity requires to be changed for negative polarity.

Figure-2: Three-Phase RV Multilevel Topology



Sources: Authors Compilation.

This topology easily extends to higher voltage levels by duplicating the middle stage as shown in Fig. 1. Therefore, this topology is modular and can be easily increased to higher voltage levels by adding the middle stage in Fig. 1. It can also be applied for three-phase applications with the same principle. This topology uses isolated dc supplies. Therefore, it does not face voltage-balancing problems due to fixed dc voltage values. In comparison with a cascade topology, it requires just one-third of isolated power supplies used in a cascade-type inverter. In Fig. 2, the complete three-phase inverter for seven levels is shown with a three-phase delta connected system. According to Fig. 2, the multilevel positive voltage is fed to the full-bridge converter to generate its polarity. Then, each full bridge converter will drive the primary of a transformer. The secondary of the transformer is delta (Δ) connected and can be connected to a three-phase system. This topology requires fewer components in comparison to conventional inverters. Another advantage of the topology is that it just requires half of the conventional carriers for SPWM controller. SPWM for seven-level conventional converters consists of six carriers, but in this topology, three carriers are sufficient. The reason is that, according to Fig. 1, the multilevel converter works only in positive polarity and does not generate negative polarities. Therefore, it implements the multilevel inverter with a reduced number of carriers, which is a great achievement for inverter control. It is also comparable to single-carrier modulation, while this topology requires the same number of signals for PWM. However, this topology needs one modulation signal which is easier to generate as opposed to the single-carrier modulation method which needs several modulation signals [29]. Another disadvantage of this topology is that all switches should be selected from fast switches, while the proposed topology does not need fast switches for the polarity generation part. In the following sections, the superiority of this topology with respect to PWM switching and number of components is discussed.

Table-1: Switching Sequences for Each Level

level / Mode	0	1	2	3
1	2,3,4	2,3,5	1,4	1,5
2		2,4,6	2,6,5	

Sources: Authors Compilation.

B. Switching Sequences

Switching sequences in this converter are easier than its counter parts. According to its inherent advantages, it does not need to generate negative pulses for negative cycle control. Thus, there is no need for extra conditions for controlling the negative voltage. Instead, the reversing full-bridge converter performs this task, and the required level is produced by the high-switching-frequency component of the inverter. Then, this level is translated to negative or positive according to output voltage requirements. This topology is redundant and flexible in the switching sequence. Different switching modes in generating the required levels for a seven-level RV inverter are shown in Table1. In Table1, the numbers show the switch according to Fig. 1 which should be turned on to generate the required voltage level. According to the table, there are six possible switching patterns to control the inverter. It shows the great redundancy of the topology. However, as the dc sources are externally adjustable sources (dc power supplies), there is no need for voltage balancing for this work. In order to avoid unwanted voltage levels du ing switching cycles, the switching modes should be selected so that the switching transitions become minimal during each mode transfer. This will also help to decrease switching power dissipation. According to the aforementioned suggestions, the sequences of switches (2–3-4), (2-3-5), (2-6-5), and (1, 5) are chosen for levels 0 up to 3, respectively. These sequences are shown in Fig. 3. As can be observed from Fig. 3, the output voltage levels are generated in this part by appropriate switching sequences. The ultimate output voltage level is the sum of voltage sources, which are included in the current path that is marked in bold. In order to produce seven levels by SPWM, three saw-tooth waveforms for carrier and a sinusoidal reference signal for modulator are required as shown in Fig. 4. In this paper, PD SPWM is adopted for its simplicity [30]. Carriers in this method do not have any coincidence, and they have definite offset from each other. They are also in phase with each other. The modulator and three carriers for SPWM are shown in Fig. 4.

Figure-3: Switching Sequence for Different Level Generation

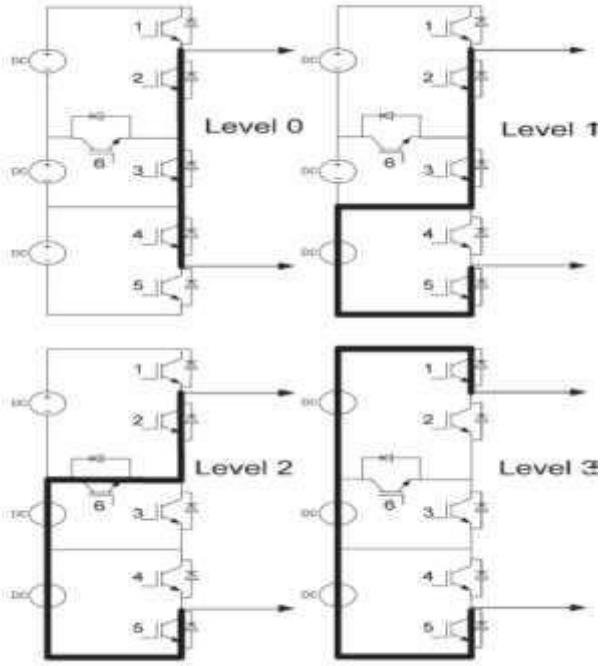
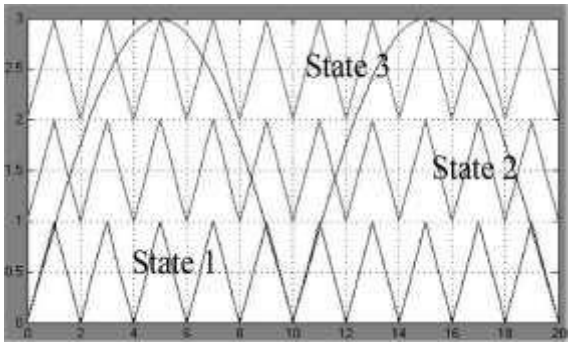


Fig. 3. Switching sequences for different level generation.

Sources: Authors Compilation.

Figure-4: SPWM Carrier and Modulator for RV Topology



Sources: Authors Compilation.

Table-2: Switching Cases in Each State According to Related Comparator Output

States	One		Two		Three	
Compare	+	-	+	-	+	-
Mode	2-3-5	2-3-4	2-5-6	2-3-5	1-5	2-5-6

Sources: Authors Compilation.

According to Fig. 4, three states are considered. The first state is when the modulator signal is within the lowest carrier. The second state is when it is within the middle carrier. Finally, the third one is when it is within the highest carrier. In each state, certain switching patterns are adopted to cover the voltage requirements. According to this definition, the switching states and switching modes are described in Table2.

Table2 shows the relation between the right comparator output according to the current state and required states for switching to meet the voltage requirements. The right comparator here refers to the comparator output of the current state. As illustrated in Table2, the transition between modes in each state requires minimum commutation of switches to improve the efficiency of the inverter during switching states. The number of switches in the path of conducting current also plays an important role in the efficiency of overall converter. For example, a seven-level cascade topology [31], [32] has 12 switches, and half of them, i.e., six switches, conduct the inverter current in each instance. However, the number of switches which conduct current in the proposed topology ranges from four switches (for generating level 3) to five switches conducting for other levels, while two of the switches are from the low-frequency (polarity generation) component of the inverter.

Therefore, the number of switches in the proposed topology that conduct the circuit current is lower than that of the cascade [33] inverter, and hence, it has a better efficiency. The same calculation is true in a topology mentioned in [34]. The least number of switches in the current path for a seven-level inverter according to [34] is five (for generating level 3), which requires one switch more in the current path compared to the proposed topology which requires only four conducting switches.

These switching sequences can be implemented by logic gates or DSP. The signal stage should be isolated from the power stage by opto couplers for control circuit protection. The drive circuit is also responsible to generate the dead time between each successive switching cycle across the dc source.

The gating signal for the output stage, which changes the polarity of the voltage, is simple. Low-frequency output stage is an H-bridge inverter and works in two modes: forward and reverse modes. In the forward mode, switches 8 and 9 as in Fig. 1 conduct, and the output voltage polarity is positive.

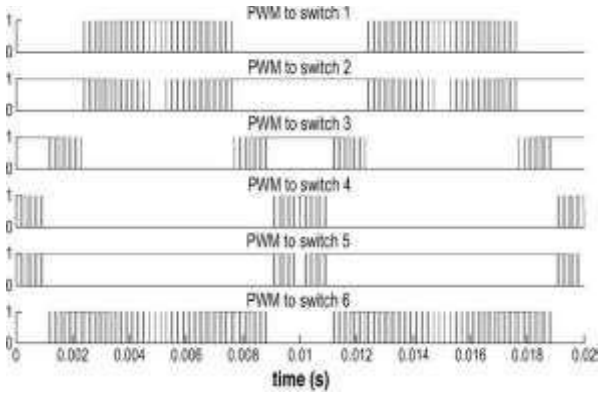
However, switches 7 and 10 conduct in reverse mode, which will lead to negative voltage polarity in the output. Thus, the low-frequency polarity generation stage only determines the output polarity and is synchronous with the line frequency.

The resulting PWM waveforms for driving the high frequency switches in the level generation part are illustrated for one complete cycle in Fig. 5. According to Fig. 5, high frequency switches can be adopted in this stage based on the required frequency and voltage level. However, low-frequency polarity generation part drive signals are generated with line frequency (50 Hz), and they only change at zero-voltage crossings.

C. Number of Components

As mentioned earlier, one of the promising advantages of topology is that it requires less high-switching-frequency components. High-frequency switches and diodes are expensive and are more prone to damages than low-frequency switches. According to MIL-HDBK-217F standard, the reliability of a system is indirectly proportional to the number of its components. Therefore, as the number of high-frequency switches is increased, reliability of converter is decreased.

Figure-5: Complete Gate Signals for Level Generation Part.



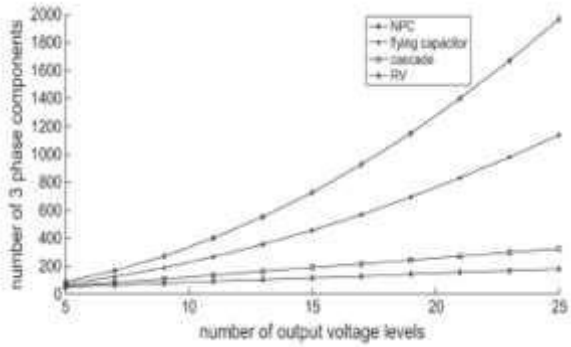
Sources: Authors Compilation.

Table-3: Number of Components for Three-Phase Inverters

Inverter type	NPC	Flying capacitor	Cascade	RV
Main switches	6(N-1)	6(N-1)	6(N-1)	3((N-1)+4)
main diodes	6(N-1)	6(N-1)	6(N-1)	3((N-1)+4)
Clamping diodes	3(N-1)(N-2)	0	0	0
DC bus capacitors/ Isolated supplies	(N-1)	(N-1)	3(N-1)/2	(N-1)/2
Flying capacitors	0	$\frac{3}{2}$(N-1)(N-2)	0	0
Total numbers	(N-1)(3N+7)	$\frac{1}{2}$(N-1)(3N+20)	$\frac{27}{2}$(N-1)	(13N+35)/2

Sources: Authors Compilation.

Figure-6: Components for Multilevel Inverters



Sources: Authors Compilation.

Figure-7: Reqired Switch for Multilevel Inverters

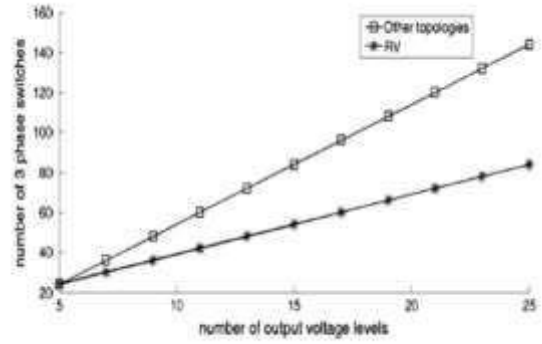


Fig. 7. Required switches for multilevel inverter.

Sources: Authors Compilation.

In the proposed converter, as can be seen, half of the switches in the full-bridge converter will not require to be switched on rapidly since they are only switched at zero crossings operating at line frequency (50 Hz). Thus, in this case, the reliability of the converter and also related expenses are highly improved. The number of required three-phase components according to output voltage levels (*N*) is illustrated in Table III [35].

It can clearly be inferred that the number of components of the proposed topology is lower than that of other topologies even more so as the voltage levels increase and it will decrease tremendously with higher voltage levels.

Fig. 6 shows the required components versus different voltage levels as mentioned in Table III. As the most important part in multilevel inverters is the power semiconductor switches which define the reliability and Fig. 7. Required switches for multilevel inverter. Control complexity, the number of required switches against the required voltage levels is shown in Fig. 7 for the new topology as well as other topologies.

According to Figs. 6 and 7, the new topology requires fewer components and also fewer switches compared to others. Therefore, it should have the potential of finding widespread applications in high-voltage power devices and apparatus that includes FACTS and HVDC. It also requires less number of components as to conventional inverters that use phase shift transformers for increasing the output voltage levels.

STATCOM, which is a type of FACTS apparatus and has been widely developed in recent years, can be a good candidate for applying the topology. In order to fulfill the stringent voltage harmonic standards such as IEEE519, a STATCOM of the conventional 48-pulse inverter is normally used [36]. The topology requires eight three-phase transformers and eight full-bridge inverters requiring 48 switches. However, the proposed topology is superior compared to this conventional topology since it requires 84 switches for implementing similar output voltage waveform with the same quality while omitting eight bulky transformers.

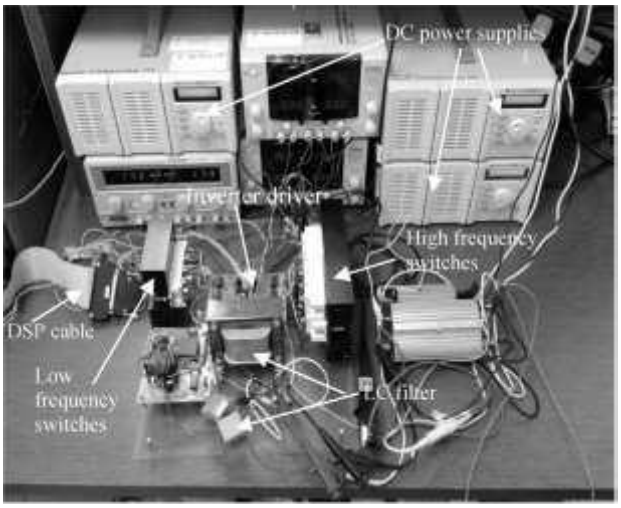
EXPERIMENTAL RESULTS

In this section, practical results are demonstrated for a single phase seven-level inverter with the proposed topology as shown in Fig. 8. The PWM controller is implemented by a dSPACE 1104 DSP card. The output *LC* filter is used to remove high-frequency switching ripples and is a combination of a 2-μF capacitor and a 10-mH inductor.

The three modules in the level generation part are SKM 50 GB, and the full-bridge module in the polarity generation part is SKM 40 GB. DC power supplies are also adjusted at 50 V. The topology is used to generate seven voltage levels for a resistive and inductive load. The output voltage is 300 VP−P.

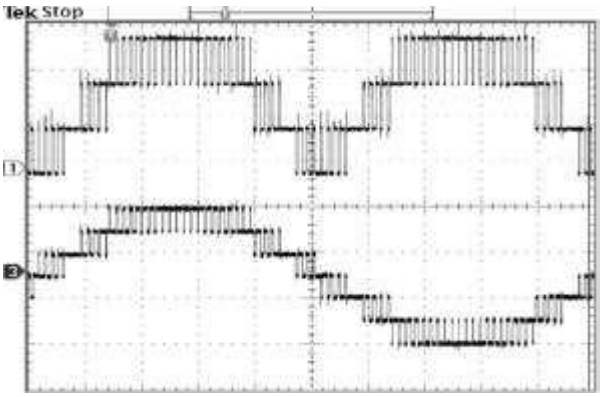
The switching frequency is 4 kHz. The PWM signals are generated for the inverter by the DSP board according to the PD-SPWM method as mentioned. The control signals are unipolar, and there is no need for generation of bipolar signals in DSP for modulation [25] in the proposed multilevel inverter, and the unipolar modulation presents better results in terms of harmonics [37]. Fig. 9 shows how the output voltage is made in the polarity generation part.

Figure-8: Experimental Prototype for the Proposed Seven -Level Topology



Sources: Authors Compilation.

Figure-9: Waveforms In the Proposed Topology from Top: Output Voltage of Polarity Generating Part (50 V/div) and output voltage (100 V/div). Time: 2 ms/div

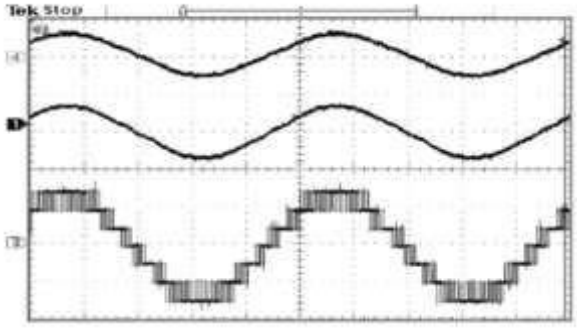


Sources: Authors Compilation.

This part produces the output polarity according to the system requirements. In Fig. 9, the voltage output of the level generating part and output of the polarity generation part are shown. According to Fig. 9, the voltage, which is generated by the high-frequency level generating component, is fed to the

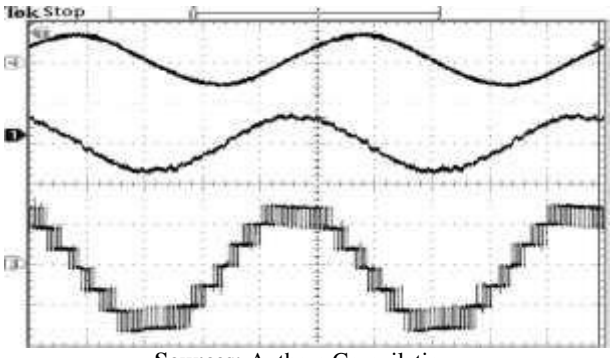
polarity generating component to define its polarity. The complete output waveform is then created, resulting in the production of the desired voltage waveform of the multilevel inverter as shown in Fig. 9. The waveform of the proposed multilevel inverter with an output filter and a resistive load of 124 Ω is shown in Fig. 10. The resulting current THD was 3.85%, which complies with the IEEE 519 harmonic standard. Figs. 10 and 11 clearly show the performance of the proposed inverter for resistive and inductive loads. Therefore, this proposed topology is proven to work with different kinds of loads as shown in Figs. 10 and 11. The output waveform with an inductive load of 182 mH, which is equal to 115 Ω, is also shown in Fig. 11. The resulting THD for current was 2.3%, which is in compliance with the IEEE 519 harmonic standard.

Figure-10: Waveform of the Proposed Multilevel Inverter With a Resistive Load from Top: Output Current (2 A/div), Output Voltage after Filter (200 V/div), and Multilevel Output Before Filter (100 V/div). Time: 4 ms/div.



Sources: Authors Compilation.

Figure-11: Waveform of the proposed multilevel inverter with inductive load from top: Output current (2 A/div), output voltage after filter (200 V/div), and multilevel output before filter (100 V/div). Time: 4 ms/div.



Sources: Authors Compilation.

CONCLUSIONS

In this paper, a new inverter topology has been proposed which has superior features over conventional topologies in terms of the required power switches and isolated dc supplies, control requirements, cost, and reliability. It is shown that this topology can be good candidate for converters used in power applications such as FACTS, HVDC, PV systems, UPS, etc. In the mentioned topology, the switching operation is separated into high- and low-frequency parts. This will add up to the efficiency of the converter as well as reducing the size and cost of the final prototype. The PD-SPWM control method is used to drive the inverter. The PWM for this topology has fewer complexities since it only generates positive carriers for PWM control. The experimental results of the developed prototype for seven-level inverter of proposed topology are demonstrated in this paper. The results clearly show that the proposed topology can effectively work as a multilevel inverter with a reduced number of carriers for PWM.

REFERENCES

[1] Ehsan, Najafi, and Abdul, Halim Mohamed Yatim, (2012). "Design and Imlementation of a New Multilevel Inverter Topology", IEEE Trans. Ind. Electronics, November, Volume: 59, Number: 11.

[2] K., Jang-Hwan; S. K. Sul, and P., N. Enjeti, (2008). "A Carrier-Based PWM Method with Optimal Switching Sequence for a Multilevel Four-Leg Voltage Source Inverter", IEEE Trans. Ind. Appl., July/August, Volume: 44, Number: 4, pp. 1239–1248.

[3] S., Srikanthan, and M., K. Mishra, (2010). "DC Capacitor Voltage Equalization in Neutral Clamped Inverters for DSTATCOM Application", IEEE Trans. Ind. Electron., August, Volume: 57, Number: 8, pp. 2768–2775.

[4] L., M. Tolbert; F., Z. Peng, and T., G. Habetler, (1999). "Multilevel Converters For Large Electric Drives", IEEE Trans. Ind. Appl., January/February, Volume: 35, Number: 1, pp. 36–44.

[5] T., L. Skvarenina, (2002). The Power Electronics Handbook, Boca Raton, FL: CRC Press.

[6] X., Yun; Y., Zou; X., Liu, and Y., He, (2007). "A Novel Composite Cascade Multilevel Converter", Proceeding 33rd IEEE IECON, pp. 1799–1804.

[7] R., H. Osman, (1999). "A Medium-Voltage Drive Utilizing Series-Cell Multilevel Topology For Outstanding Power Quality", Conference Rec. 34th IEEE IAS Annuual Meeting, Volume: 4, pp. 2662–2669.

[8] E., Najafi, and A., H. M. Yatim, (2011). "A Novel Current Mode Controller for a Static Compensator Utilizing Goertzel Algorithm to Mitigate Voltage Sags", Energy Convers. Manage., April, Volume: 52, Number: 4, pp. 1999–2008.

[9] N., Seki, and H., Uchino, (1997). "Converter Configurations And Switching Frequency for a GTO Reactive Power Compensator", IEEE Trans. Ind. Appl., July/August, Volume; 33, Number: 4, pp. 1011–1018.

[10] G., Shahgholiyan; E., Haghjou, and S., Abazari, (2009). "Improving the Mitigation of Voltage Flicker by Usage of Fuzzy Control in a Distribution Static Synchronous Compensator (DSTATCOM)", Majlesi Journal Elect. Eng., June, Volume: 3, Number: 2, pp. 25–35.

[11] K., Nakata; K., Nakamura; S., Ito, and K., Jinbo, (1994). "A Three-Level Traction Inverter with IGBTs for EMU," in Conference Rec. IEEE IAS Annual Meeting, Volume: 1, pp. 667–672.

[12] A., Jidin; N., R. N. Idris; A., H. M. Yatim; T., Sutikno, and M., E. Elbuluk, (2011). "An Optimized Switching Strategy For Quick Dynamic Torque Control In DTC-Hysteresis-Based Induction Machines", IEEE Trans. Ind. Electron, August, Volume: 8, Number: 8, pp. 3391–3400.

[13] K., Y. Lau; M., F. M. Yousof; S., N. M. Arshad; M., Anwari, and A., H. M. Yatim, (2010). "Performance Analysis of Hybrid Photovoltaic/Dieselenergy System under Malaysian Conditions", Journal Energy, August, Volume: 35, Number: 8, pp. 3245–3255.

[14] G., M. Martins; J., A. Pomilio; S., Buso, and G., Spiazzi, (2006). "Three-Phase Lowfrequency Commutation Inverter for Renewable Energy Systems", IEEE Trans. Ind. Electron, October, Volume: 53, Number: 5, pp. 1522–1528.

[15] S., Daher; J., Schmid, and F., L. M. Antunes, (2008). "Multilevel Inverter Topologies for Stand-Alone PV Systems", IEEE Trans. Ind. Electron, July, Volume; 55, Number: 7, pp. 2703–2712.

[16] R., Teodorescu; F., Blaabjerg; J., K. Pedersen; E., Cengelci, and P., N. Enjeti, (2002). "Multilevel Inverter by Cascading Industrial VSI", IEEE Trans. Ind. Electron, August, Volume: 49, Number: 4, pp. 832–838.

[17] D., A. B. Zambra; C., Rech, and J., R. Pinheiro, (2007). "A Comparative Analysis between the Symmetric and the Hybrid Asymmetric Nine-Level Series Connected H-Bridge Cells Inverter", Proceeding Eur. Conf. Power Electron. Appl., pp. 1–10.

[18] G., S. Perantzakis; F., H. Xepapas, and S., N. Manias, (2007). "A Novel Four-Level Voltage Source Inverter-Influence of Switching Strategies on the Distribution of Power Losses", IEEE Trans. Power Electron, January, Volume 22, Number: 1, pp. 149–159.

[19] E., Babaei, (2010). "Optimal Topologies for Cascaded Sub-Multilevel Converters", Journal Power Electron, May, Volume: 10, Number: 3, pp. 251–261.

[20] G., Mondal; K., Gopakumar; P., N. Tekwani, and E., Levi, (2007). "A Reduced Switch - Count Five-Level Inverter With Common-Mode Voltage Elimination for An Open-End Winding Induction Motor Drive", IEEE Trans. Ind. Electron, August, Volume: 54, Number: 4, pp. 2344–2351.

[21] E., Beser; B., Arifoglu; S., Camur, and E., K. Beser, (2010). "Design and Application of a Single Phase Multilevel Inverter Suitable for Using as a Voltage Harmonic Source", J. Power Electron, March, Volume: 10, Number: 2, pp. 138–145.

[22] G., Ceglia; V., Guzman; C., Sanchez; F., Ibanez; J., Walter, and M., I. Gimenez, (2006). "A New Simplified Multilevel Inverter Topology for DC-AC Conversion", IEEE Trans. Power Electron., September, Volume: 21, Number: 5, pp. 1311–1319.

[23] N., A. Rahim; K., Chaniago, and J., Selvaraj, (2011). "Single-phase Seven-Level Grid-Connected Inverter for Photovoltaic System", IEEE Trans. Ind. Electron., June, Volume: 58, Number; 6, pp. 2435–2443.

[24] J., Selvaraj, and N., A. Rahim, (2009). "Multilevel Inverter for Grid-Connected PV System Employing Digital PI Controller", IEEE Trans. Ind. Electron., January, Volume; 56, Number: 1, pp. 149–158.

[25] S., G. Song; F., S. Kang, and S.,-J. Park, (2009), "Cascaded Multilevel Inverter Employing Three-Phase Transformers and Single DC Input", IEEE Trans.Ind. Electron, June, Volume 56, Number: 6, pp. 2005–2014.

[26] P., Barbosa; P., Steimer; J., Steinke; L., Meysenc; M., Winkelnkemper, and N., Celanovic, (2005). "Active Neutral-Point-Clamped Multilevel Converters", Proceeding IEEE 36th Power Electron Spec. Conference, pp. 2296–2301.

[27] S., A. Gonzalez; M., I. Valla, and C., F. Christiansen, (2007). "Analysis of a Cascade Asymmetric Topology Formultilevel Converters", Proceeding. IEEE ISIE, pp. 1027–1032.

[28] P., Lezana, and J., Rodriguez, (2007). "Mixed Multicell Cascaded Multilevel Inverter", Proceding IEEE ISIE, 2007, pp. 509–514.

[29] R., Stala, (2011). "Application of Balancing Circuit for DC-Link Voltages Balance In a Single-Phase Diode-Clamped Inverter with Two Three-Level Legs", IEEE Trans. Ind. Electron., September, Volume; 58, Number: 9, pp. 4185–4195.

[30] C., Govindaraju, and K., Baskaran, (2010). "Analysis and Implementation of Multiphase Multilevel Hybrid Single Carrier Sinusoidal Modulation", J. Power Electron., July, Volume: 10, Number: 4, pp. 365–373.

[31] A., Radan; A., H. Shahirinia, and M., Falahi, (2007). "Evaluation of Carrier-Based PWM Methods for Multi-Level Inverters", Proceeding IEEE ISIE, pp. 389–394.

[32] Y.,-M. Park; H.,-S. Ryu; H.,-W. Lee; M.,-G. Jung, and S.,-H. Lee, (2010). "Design of a Cascaded H-bridge Multilevel Inverter Based on Power Electronics Building Blocks and Control for High Performance", Journal Power Electron., May, Volume: 10, Number: 3, pp. 262–269.

[33] M., Malinowski; K., Gopakumar; J., Rodriguez, and M., A. Pe´rez, (2010). "A Survey on Cascaded Multilevel Inverters", IEEE Trans. Ind. Electron., July, Volume: 57, Number: 7, pp. 2197–2206.

[34] C. Cecati, F. Ciancetta, and P. Siano, "A multilevel inverter for photovoltaic systems with fuzzy logic control," IEEE Trans. Ind. Electron., vol. 57, no. 12, pp. 4115–4125, Dec. 2010.

[35] G.,-J. Su, (2005). "Multilevel DC-Link Inverter", IEEE Trans. Ind. Appl., May/June, Volume: 41, Number: 3, pp. 848–854.

[36] J.,-S. Lai, and F., Z. Peng, (1996). "Multilevel Converters-a New Breed of Power Converters", IEEE Trans. Ind. Appl., May/June, Volume: 32, Number: 3, pp. 509–517.

[37] C., A. C. Cavaliere; E., H. Watanabe, and M., Aredes, (2002). "Multi-pulse STATCOM Operation under Unbalanced Voltages", Proceeding IEEE Power Eng. Soc. Winter Meeting, Volume: 1, pp. 567–572.

[38] S., Halasz; G., Csonka; A., A. M. Hassan, and B., T. Huu, (1996). "Analysis of the Unipolar PWM Techniques", Proc. 8th MELECON, Volume: 1, pp. 353–356.

*****

CHARACTERISATION OF WELDMENTS OF NITROGEN ALLOYED Cr-Mo STEELS

I. Esther[87] N. Murugan[88]

ABSTRACT

Nitrogen alloyed Cr-Mo steels being a special material are finding increasing applications in energy, chemical, petrochemical, fertilizer, oil, gas, mining and transportation sectors due to their outstanding combination of mechanical properties as well as corrosion, oxidation and wear resistance. The AS cast nitrogen alloyed Cr-Mo steel bar was autogenously welded with GTAW process at 3 different heat input. The welded samples were analyzed to find the composition of carbon and nitrogen and microstructural analysis were carried out. Finally corrosion rate and hardness value of the weld metal were found for different heat input condition. In weld metal at low heat input condition less amount of nitrogen escaped from weld metal, which enhanced corrosion resistance of the weld metal. It was observed that at lower heat input the rate of solidification was high which resulted in finer grain size and increased the hardness of weld metal.

KEYWORDS

Steel, Cr-Mo, Nitrogen Alloyed, Corrosion, Metal, Weld, etc.

INTRODUCTION

The nitrogen solubility in nitrogen alloyed steel increases with alloying elements like Mo, Mn, Cr, V, Nb, Ti and the solubility decreases with alloying elements like Ni, Si and C. Since the nitrogen pick up from the gas phase is limited by its solubility given by Sievert's law, suitable melting techniques as well as choice of alloying elements are essential to retain the required quantity of nitrogen content under normal atmospheric pressure. One of the challenges involved in Nitrogen alloyed Cr-Mo Steel production is to obtain a fully dense component with uniform nitrogen content and excellent surface properties. Uniform nitrogen distribution leads to more homogeneous microstructure, giving better mechanical properties.

The nitrogen concentration remaining in the solidified weld metal represents a balance between nitrogen absorption, which is dominated by one of the faster processes-atomic ionic or NO gas and the evolution of nitrogen by bubble formation. In welding of nitrogen containing stainless steels, it is essential to avoid nitrogen losses, which could result in loss of corrosion resistance as well as mechanical properties. Welding of nitrogen containing stainless steels has been performed for many years with good results.

Mohamed Nazirudeen et al. reported that the solubility of nitrogen content in martensitic stainless steel could be varied depending on the percentage of Cr, Mo and V content in conventional melting route (Mohamed Nazirudeen et al., 2003).

[87] Assistant Professor, Coimbatore Institute of Technology, Tamil Nadu, India, arulestherjo@gmail.com
[88]Professor, Coimbatore Institute of Technology, Tamil Nadu, India, murugan@cit.edu.in

Welding parameters such as shielding gases, welding speed, arc current and arc voltage etc, are very important variables affecting nitrogen absorption in weld metal. The nature of gases in the welding atmosphere plays a vital role in the nitrogen absorption. Oxidizing gases such as O_2 and CO_2 increase nitrogen absorption, whereas a reducing gas such as hydrogen decreases the amount of nitrogen absorbed into the pool (Kobayashi et al., 2002). Nitrogen content of up to 0.4% can be attained in the weld metal depending on the nature of shielding gases and the partial pressures. However, limited literatures are available on nitrogen alloyed Cr-Mo steels and their weldability. It is essential to explore the effects of welding process parameters on mechanical, metallurgical and corrosion properties of nitrogen alloyed Cr Mo steels.

EXPERIMENTAL PROCEDURE

The present work was carried out in two phases. First phase dealt with characterization of as casted nitrogen alloyed Cr Mo steel. The second phase involves GTA welding and characterization of nitrogen alloyed Cr Mo welded steels.

Characterisation of AS Cast Nitrogen Alloyed Cr - Mo steels

The chemical composition of the as cast nitrogen alloyed Cr Mo steel plates were analyzed by ARL 2460 spectrovac system. The microstructure of the nitrogen alloyed Cr Mo steel was analysed using the standard metallurgical procedure. Its corrosion rate was measured using AC-GILL Potentiostat and Galvanostat as per standard procedure and was found to be test 1.99 mdd. The microhardness values were measured using Zwick Microhardness machine.

Welding of Nitrogen Alloyed Cr-Mo Steel

The plates of size 35 mm X 35 mm X 3.5 mm were machined from the as cast nitrogen alloyed Cr-Mo steel bar. It was autogenously welded with GTAW process at three different heat input conditions obtained by changing welding speed from 1.66 mm/s to 3.33 mm/ sec, welding current from 70 A to 100 A and arc voltage from 13 V to 15 V. A PANA make -TIG - 300 welding machine (with pulsed setup) was used in this investigation.

Composition Analysis

Spectro vac ARL-2460 Iron Base analyzer was used to analyse the composition of the weld metal especially carbon and nitrogen present in the welded plates. The changes in the heat input changes the Composition mainly the percentage of Carbon and Nitrogen in the welded specimen. At high heat input condition the evolution of Carbon and Nitrogen to the atmosphere was increased**.** Due to this reason solidification cracking and blowholes were present in the weld metal. At low heat input condition, the nitrogen evolution was reduced which tends to increase its corrosion resistance.

Microstructural Studies

The prepared samples for metallographic analysis were finally polished on velvet cloth using diamond paste and kerosene.

The polished samples were cleaned with running water, swabbed with cotton and dried. Then it was etched with 2% Nital to get core microstructure. The samples were observed under Optical Microscope and Photomicrographs were taken using Nikon Epiphot Microscope.

Corrosion Studies

The corrosion rates of the welded specimens were obtained from Electrochemical Potentiokinetic Reactivation (EPR) test conducted with an AC GILL Potentiostat system. The corrosion study was done in 10 % Hydrochloric acid (HCL). The system was designed so that only an extremely small current can pass between the reference electrodes and working electrode; the current needed to polarize the working electrode was supplied from the auxiliary electrode. Potentiostat actually contains an operational amplifier or similar device in order to provide the feedback of the current necessary to maintain the fixed potential between the working and reference electrode.

Hardness Testing

Hardness test was carried out at 10-kg load in Zwick Microhardness Tester. The hardness was checked at three different locations and the average value was reported and compared at different heat input conditions.

RESULTS AND DISCUSSIONS

Case-1: Characterisation of As Cast Nitrogen Alloyed Cr - Mo Steels

The chemical composition of AS cast nitrogen alloyed Cr Mo steel plates were analyzed by ARL 2460 spectrovac system. The % by weight composition is presented in Table-1 below:

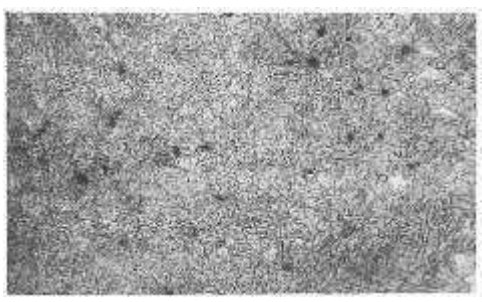
Table-1: Chemical Composition of Nitrogen Alloyed Chromium Steels

S. No.	C %	Si %	S %	P %	Mn %	Ni %	Cr %	Mo %	V %	Cu %	N %
M.1	0.096	0.36	0.02	0.02	0.47	0.13	8.22	1.08	0.22	0.03	0.05

Sources: Authors Compilation.

The microstructure of nitrogen alloyed Cr Mo steel was analysed using the standard metallurgical procedure and a typical microstructure is shown in Fig.1. Its corrosion rate was measured using AC-GILL Potentiostat and Galvanostat as per standard procedure and was found to be test 1.99 mdd. Microhardness values were measured using Zwick Microhardness machine and average value found to be 460 HV.

Figure-1: Typical Microstructure of the Base Metal (Magnification 200 X)



Sources: Authors Compilation.

Case-2 Welding Of Nitrogen Alloyed Cr-Mo Steel

The plates of size 35 mm X 35 mm X 3.5 mm were machined from the as cast nitrogen alloyed Cr-Mo steel bar. It was autogenously welded with GTAW process at three different heat input conditions obtained by changing welding speed from 1.66 mm/s to 3.33 mm/ sec, welding current from 70 A to 100 A and arc voltage from 13 V to 15 V. PANA-TIG wp-300 (with pulse) machine was used.

Composition Analysis

Spectro vac ARL-2460 Iron Base analyzer was used to analyse the composition of the weld metal especially carbon and nitrogen present in the welded plates. Table 3 gives the composition of carbon and nitrogen in the weld metal. The changes in the heat input changes the Composition mainly the percentage of Carbon and Nitrogen in the welded specimen.

At high heat input condition the evolution of Carbon and Nitrogen to the atmosphere was increased**.** Due to this reason solidification cracking and blowholes were present in the weld metal. At low heat input condition, the nitrogen evolution was reduced which tends to increase its corrosion resistance. The percentage of Carbon and Nitrogen are shown in Table.2.

Table-2: Heat Input and Percentage of Carbon and Nitrogen in Weld Metal

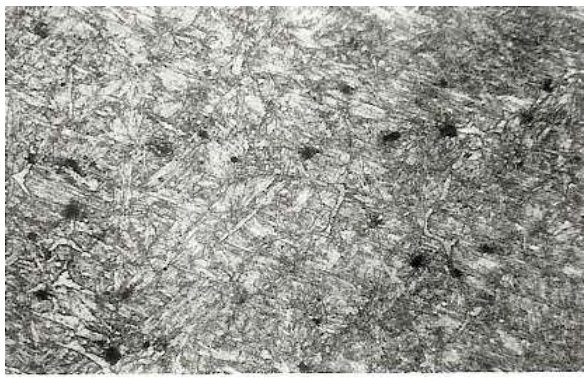
S. No	Welding speed mm/s	Current Amp	Voltage Volts	Heat Input J/mm	Carbon % By wt	Nitrogen % By wt
1.	1.66	100	14	843.37	0.081	0.01
2.	2.5	90	14	504	0.086	0.015
3.	3.33	80	13	312	0.092	0.025

Sources: Authors Compilation.

Microstructural Studies

The prepared samples for metallographic analysis were finally polished on velvet cloth using diamond paste and kerosene. The polished samples were cleaned with running water, swabbed with cotton and dried. Then it was etched with 2% Nital to get core microstructure. Then, samples were observed under Optical Microscope and Photomicrographs were taken using Nikon Epiphot microscope equipment. Fig. 2 show typical microstructure of the weld metal.

Figure-2: Typical Microstructure of the weld metal (Heat input 843.37 J / mm)



(Magnification 200 x)
Sources: Authors Compilation.

Corrosion Studies

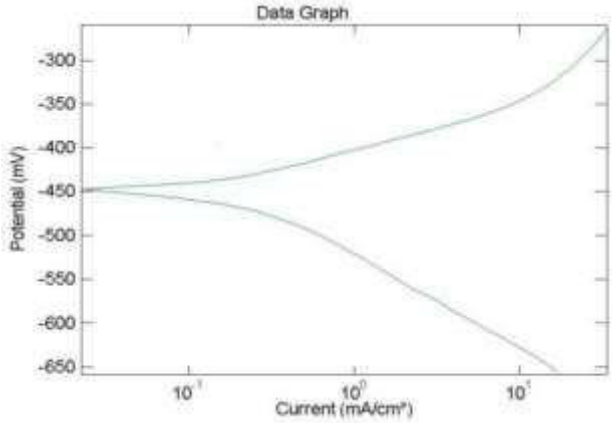
The corrosion rate of the welded specimens were obtained from electrochemical polarization test conducted using AC GILL Potentiostat and Galvanostat system. The corrosion study was done in 10% Hydrochloric acid (HCL). The system was designed so that only an extremely small current can pass between the reference electrodes and working electrode; the current needed to polarize the working electrode was supplied from the auxiliary electrode. The potentiostat actually contains an operational amplifier or similar device in order to provide the feedback of the current necessary to maintain the fixed potential between the working and reference electrode. By plotting potential verses log Iapp, a linear relationship between current density and potential was developed. Fig3 shows a typical Toffel Curve of weld metal specimen. The estimated Corrosion rate at different heat input condition are shown in Table 3.

Table-3: Corrosion Rate of Weld Metal

S. No.	Heat Input J/mm	Corrosion Rate mdd
1.	843.37	4.214
2.	504	3.657
3.	312	2.382

Sources: Authors Compilation.

Figure-3: Typical Toffel Curve of Weld Metal Specimen in Chloride Solution



Sources: Authors Compilation.

Hardness Testing

Hardness test was carried out at 10-kg load in Zwick microhardness machine. Hardness was checked at three different locations and the average value was reported and compared at different heat input conditions. The average hardness values at different heat input condition are shown in Table4.

Table-4: Corrosion Rate of Weld Metal

S. No.	Heat Input J/mm	Hardness in HV
1.	843.37	262
2.	504	341
3.	312	453

Sources: Authors Compilation.

CONCLUSIONS

GTA welding of nitrogen alloyed Cr-Mo steel was carried out. Its mechanical and metallurgical characterization of both base metal and weld metal were performed. A few of conclusions arrived at are given below.

As Cast Sample

1. As Cast samples reveal non-uniform microstructure with lower hardness.
2. Corrosion rate is very low in as cast sample indicating its high resistance to corrosion.

Welded Samples

1. When more amounts of Nitrogen and Carbon present in the base metal the corrosion resistance of base metal is high.
2. At high heat input condition, more evolution of nitrogen to the atmosphere was observed resulting in solidification cracking and blowholes in the weld metal.
3. At low heat input condition less amount of nitrogen escaped from weld metal, which enhanced corrosion resistance of the weld metal.
4. When heat input is low the rate of solidification is high resulting in finer grains size of weld metal which increases its hardness.

REFERENCES

[1] Anthamatten, (1990). "High Nitrogen 9-12% Chromium Steels for high Temperature Application", High Nitrogen Steels, HNS 90, pp. 436- 441.

[2] Berns, H., and Krafft, F., (1990). "Influence of a Thermo Mechanical Treatment on the Mechanical Properties of Creep Resistant, Steels Alloyed with Nitrogen under Pressure", High Nitrogen Steels, HNS-88, Institute of metals, London, pp. 338-342.

[3] Insu, WOO, and Yasushi, KIKUCHI, (2002), "Weldability of High Nitrogen Stainless Steel", ISIJ International, Volume: 2, Number: 12, pp. 1334 – 1343.

[4] Kobayashi, T.; Kuwana, T.; Kikuchi, Y., and J., Jap., (1971). Welding Soc., Volume: 40, pp. 221.

[5] Lenel, UK, and Knott, BR, (1987) "Structure and Properties of Corrosion and Wear Resistance Cr-Mn-N steels", Metal Trans, Volume: 18 A, Number: 5, pp. 847- 855.

[6] M., O. Speidel, (1998). "Properties and Application of High Nitrogen Steels, High Nitrogen Steels, HNS –88", Proceedings of the Inter Conference on may 1998.

[7] S., S. Mohamed Nazirudeen et al., (2003) "Effect of Alloying elements on Nitrogen Solubility in Austenitic in Stainless steels", The Proceedings of the National Conference on Quality Control in Metallurgical Industries, pp: 234-241.

*****

A HYBRID CASCADE CONVERTER TOPOLOGY WITH ADJACENT SWITCHING TECHNIQUE

D. Raja[89] S. Dhathul Farina[90]

ABSTRACT

A novel H-bridge multilevel pulsewidth modulation converter topology based on a series connection of a high-voltage diode-clamped inverter and a low-voltage conventional inverter is proposed in this paper. A dc link voltage arrangement for the new hybrid and asymmetric solution is presented to have a max-imum number of output voltage levels by preserving the adjacent switching vectors between voltage levels. Hence, a 15-level hybrid converter can be attained with a minimum number of power com-ponents. A comparative study has been carried out to present high performance of the proposed configuration to approach a very low total harmonic distortion of voltage and current, which leads to the possible elimination of the output filter. Regarding the proposed configuration, a new cascade inverter is verified by cascading an asymmetrical diode-clamped inverter, in which 19 levels can be synthesized in output voltage with the same number of components. To balance the dc link capacitor voltages for the maximum output voltage resolution as well as synthesize asymmetrical dc link combination, a new multi-output boost converter is utilized at the dc link voltage of a seven-level H-bridge diode-clamped inverter. Simulation and hardware results based on different modulations are presented to confirm the validity of the proposed approach to achieve a high-quality output voltage.

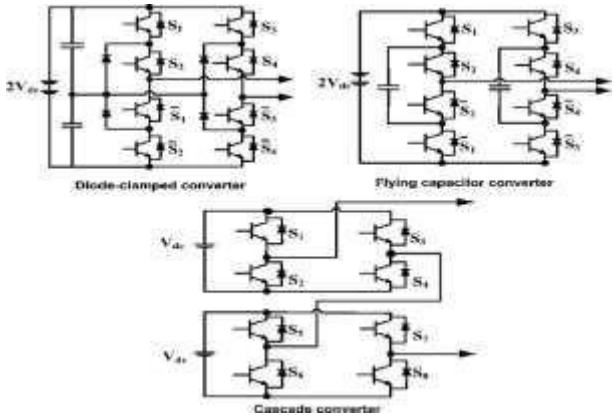
KEYWORDS

Asymmetrical Diode-Clamped Inverter, Hybrid Cascade Converter, Predictive Current Control etc.

INTRODUCTION

Multilevel power conversion has achieved wide acceptance for its capability of high-voltage (HV) and high-efficiency operation. The most popular advantages of the multilevel inverter compared with the traditional voltage source inverter are high-power-quality waveforms with lower distor-tion and a low blocking voltage by switching devices. As the number of levels increases, these advantages will be enhanced; however, it can impose a significant expense of the increase in circuit complexity, which reduces the reliability and effi-ciency in such a converter. Three prominent five-level inverter topologies, i.e., diode-clamped [1], flying capacitor [2], and cascade [3], are shown in Fig. 1.

Figure-1



Sources: Authors Compilation.

An extensive comparison between the multilevel topologies have been performed in [4] in terms of their applications, circuit modeling, modulation tech-niques, and technical issues. Among these topologies, cascade configuration has been utilized for medium-voltage and HV re-newable energy systems such as photovoltaic due to its modular and simple structure. Application of the cascade inverter for renewable energy systems is reviewed in [5] and [6].

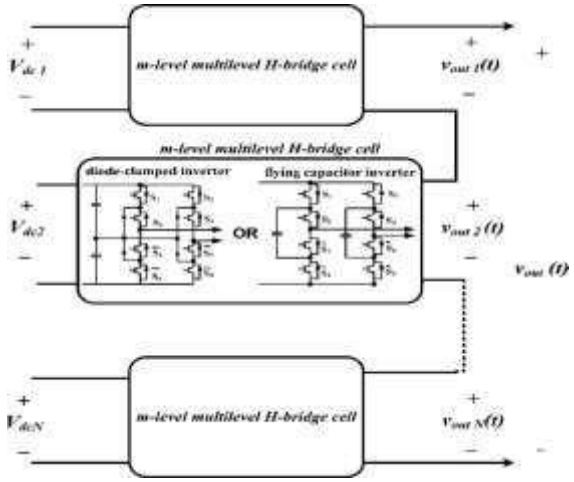
Higher level can be implemented easily by adding classical H-bridge cells to this configuration. However, it needs additional dc voltage sources and switching devices. Proposition of a cascade converter using a single dc source and capacitors is proposed in [7], which can save the extra dc sources for higher level con-verters; however, a capacitor voltage balancing algorithm is re-quired [8]. Typically, different types of multilevel converters are utilized with the same rating of the dc link voltages and power devices due to modularity and simplicity of the control strat-egy. Recently, asymmetrical multilevel inverters with unequal dc source voltages have been addressed in literature [9]–[30].

Therefore, based on different switching states, it is possible to achieve more voltage levels on output voltage by adding and subtracting dc link voltages compared with conventional mul-tilevel inverters with the same number of components [9]. By doing so, output voltage with superior quality can be obtained with less circuit and control complexity, and also, increasing the harmonic characteristic of the output voltage can decrease the size of the filter. The hybrid converters have been the main fo-cus of the literature with regard to asymmetrical configuration of multilevel inverters as they have shown their abilities and strengths in medium- and high-power applications [9], [12]. Diverse topologies have been studied based on a variety of H-bridge cascaded cells and dc voltage ratio to enhance the out-put voltage resolution compared with the same dc voltage ratio of the cells [12]–[30]. However, due to the different voltage rates of switching devices in hybrid

[89] PG Scholar, Raja College of Engineering & Technology, Tamil Nadu, India, drajaengg@gmail.com
[90] Assistant Professor, Raja College of Engineering & Technology, Tamil Nadu, India, dhafar@gmail.com
Note: Manuscript received March 7, 2009; revised May 21, 2009 and July 23, 2009. Date of current version December 27, 2010. This work was supported by Australian Research Council (ARC) under ARC Discovery Grant DP0774497. Recommended for publication by Associate Editor - B. Wang.
A. Nami, F. Zare, and A. Ghosh are with the School of Electrical Engineering, Queensland University of Technology, Brisbane, Qld. 4001, Australia (e-mail: namia@qut.edu.au; f.zare@qut.edu.au; a.ghosh@qut.edu.au).
F. Blaabjerg is with Institute of Energy Technology, Aalborg University, DK-9220 Aalborg, Denmark (e-mail: fbl@iet.aau.dk).
Digital Object Identifier 10.1109/TPEL.2009.2031115

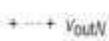
configuration, it loses its mod-ularity compared with symmetrical cascade inverters. Various pulsewidth modulation (PWM) strategies for symmetrical cas-cade inverters with high and fundamental switching frequency have been presented [14]–[16]. To reduce switching losses and improve the converter efficiency, hybrid modulations for cas-cade converters with unequal dc sources are proposed, which allows use of the slow switching device in the higher voltage cells and fast switching devices in lower voltage cells [17]–[19].

Figure-2: Topology of Cascaded Multilevel H-Bridge



Converter
Sources: Authors Compilation.

Since more voltage levels correspond to the increasing num-ber of components, recent research in this area has focused on a series of connected multilevel converters in a cascaded H-bridge structure [28]–[30]. The structure of the cascaded multilevel in-verter is demonstrated in Fig. 2, where the configuration con-sists of *m*-level multilevel H-bridge cells (either diode-clamped or flying capacitor inverter), each with an isolated dc source. The number of cells depends on the desired output volt-age level, which is synthesized by adding up all the H-bridge cells output



voltage as $v_{out}(t) = v_{out1}(t) + v_{out2}(t) + v_{out3}(t)$ (t). In a system that utilizes equal sets of dc

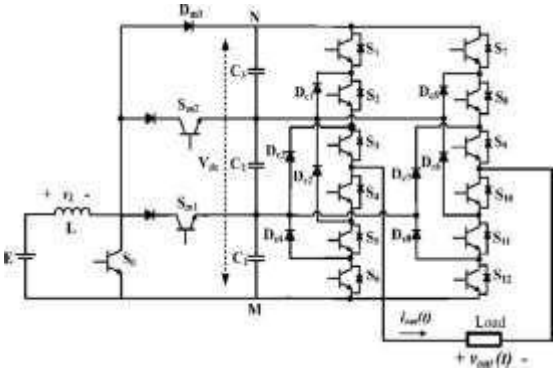
sources ($V_{dc1} = V_{dc2} = V_{dc3} = \cdots = V_{dcN}$), the number of output voltage levels is $N\times (m-1) + 1$, where N is the number of cascaded cells and m is the number of output voltage levels in each multilevel H-bridge cells. The main advantage of this ar-rangement is the simplicity to cascade several H-bridge cells for improvement of the output voltage resolution with reduced num-ber of components. However, capacitor voltage imbalance and complexity of the system can cause a critical problem that should be taken into account in this configuration using either diode-clamped or flying capacitor topology [31], [35]. To address this limitation, isolated dc sources or, alternatively, auxiliary con-verters can be used for capacitor voltage balancing. Utilization of unequal dc sources on each series diode-clamped or flying capacitor cells can increase the number of output voltage for a given power circuit in Fig.2, with the equivalent number of components. A different dc voltage ratio for H-bridge cells is proposed to achieve the maximum number of output voltage lev-els. However, along with possible maximization of obtainable output voltage levels based on the voltage ratio of dc sources, the existence of adjacent switching vectors to move from one possi-ble voltage level to another with only one switch change should be considered. Simultaneous switching of different switches is not an immense problem when there are just a few of them hap-pening over one cycle; however, when switching between the nonadjacent switching vectors occurs frequently in modulation between adjacent levels, it becomes a critical issue to increase the switching losses.

In this paper, a general idea of cascading multilevel H-bridge cells is used to propose different configurations using a seven-level symmetrical and asymmetrical diode-clamped H-bridge converter supplied with a multi-output boost (MOB) converter, cascaded with classical three-level inverters. The MOB con-verter can solve the capacitor voltage imbalance problem as well as boost the low output voltage of renewable energy sys-tems such as solar cells to the desired value of the diode-clamped dc link voltage. DC voltage ratio of cells will be presented to obtain maximum voltage levels on output voltage with adjacent switching vectors between all possible voltage levels, which can minimize the switching losses. Using a triple-output dc–dc converter offers asymmetrical dc link capacitor voltage arrange-ment for the seven-level H-bridge diode-clamped converter, in which nine voltage levels can be obtained with the same num-ber of components. Using an asymmetrical diode-clamped con-verter in proposed cascaded H-bridge cells achieves four more voltage levels in the output compared with the symmet-rical configuration. Performance of the proposed asymmetrical H-bridge diode-clamped inverter has been verified by simula-tion and hardware results. Finally, two different PWM methods based on predictive current control have been presented to vali-date the proposed approach.

II. SYMMETRICAL AND ASYMMETRICAL DIODECLAMPED CONVERTERS (SDCC AND ADCC) USING MOB CONVERTER

A new dc–dc boost converter with multiple outputs, which can be used as a front-end converter to boost the inverter's dc link voltage for grid connection systems based on a diode-clamped converter, is analyzed in [35]. Using this MOB converter, the dc voltage across each capacitor can be adjusted to a desired voltage level, thereby solving the main problem associated with balancing the capacitors' voltages in such converters.

Figure-3: H-Bridge Proposed Diode-Clamped Inverter with 3 Dc Link Capacitors Joint with Triple-Output Boost Converter



Sources: Authors Compilation.

Table-1: Possible Switching States in One Leg of a Seven-Level Diode-Clamped Inverter

Switching function States	S_1	S_2	S_3
0	0	0	0
1	0	0	1
2	0	1	1
3	1	1	1

Sources: Authors Compilation.

In [11], a new modulation technique has been presented for a diode-clamped inverter when voltages across capacitors are unequal. Fig. 3 shows a configuration for an H-bridge seven-level diode-clamped inverter joint with the front-end MOB converter. An unequal dc link arrangement is applied instead of identical dc link capacitor voltages. The bottom capacitor's voltage is kept at twice the level of other capacitors during operation, so that the configuration has asymmetrical behavior with respect to the neutral point ($V_{c1} = 2V_{c2} = 2V_{c3}$).

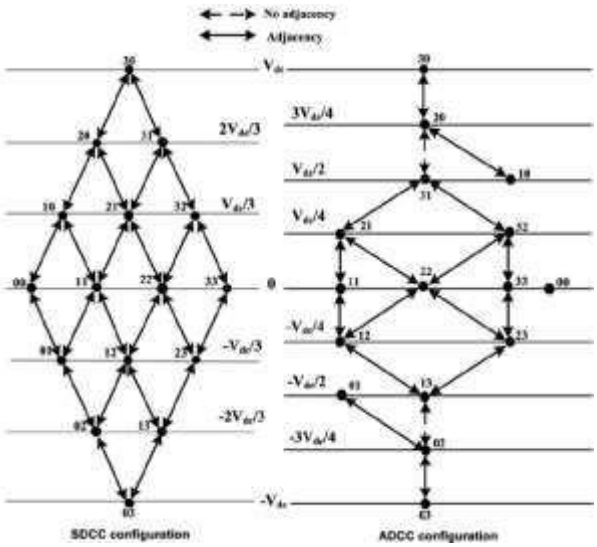
As shown in Fig. 3, it is sup-posed that the low input voltage (E) is boosted to V_{dc} at the dc link of the seven-level diode-clamped H-bridge inverter (V_{NM}). Therefore, in SDCC configuration, dc link capacitors voltage ratios are $V_{c1} = V_{c2} = V_{c3} = V_{dc}/3$; however, in ADCC con-figuration, capacitors' voltages are maintained at $V_{c1} = V_{dc}/2$ and $V_{c3} = V_{c2} = V_{dc}/4$ (with respect to the neutral point). According to the structure of the seven-level diode-clamped converter, there are four possible switching states in each leg of the inverter that can be derived from four switch combina-tions to obtain different dc link voltage levels. The "ON" and "OFF" switching states of each switch are defined as "1" and "0," respectively. Four switching states in one leg of the diode-clamped H-bridge inverter are distinguished by four switching function states that are summarized in Table I. For example, (011) means that $S_1 = 0$ (OFF), $S_2 = 1$ (ON), and $S_3 = 1$ (ON), which is defined as switching function state "2." All possible switching states associated with different output voltage levels in SDCC and ADCC are shown in Table II. Exploring the output voltage levels, nine different voltage levels can be generated in asymmetrical dc link arrangement based on different switching states, in which two more voltage lev-els can be synthesized in the output voltage compared with symmetrical arrangement with same number of switching devices.

Table-2: Output Voltage Levels Associated With Different Switching Unction States for SDCC and ADCC Configurations

Switching function states	S_1	S_2	S_3	S_4	S_5	S_6	SDCC (V_{out})	ADCC (V_{out})
00	0	0	0	0	0	0	0	0
01	0	0	0	0	0	1	$-V_{dc}/3$	$-V_{dc}/2$
02	0	0	0	0	1	1	$-2V_{dc}/3$	$-3V_{dc}/4$
03	0	0	0	1	1	1	$-V_{dc}$	$-V_{dc}$
10	0	0	1	0	0	0	$V_{dc}/3$	$V_{dc}/2$
11	0	0	1	0	0	1	0	0
12	0	0	1	0	1	1	$-V_{dc}/3$	$-V_{dc}/4$
13	0	0	1	1	1	1	$-2V_{dc}/3$	$-V_{dc}/2$
20	0	1	1	0	0	0	$2V_{dc}/3$	$3V_{dc}/4$
21	0	1	1	0	0	1	$V_{dc}/3$	$V_{dc}/4$
22	0	1	1	0	1	1	0	0
23	0	1	1	1	1	1	$-V_{dc}/3$	$-V_{dc}/4$
30	1	1	1	0	0	0	V_{dc}	V_{dc}
31	1	1	1	0	0	1	$2V_{dc}/3$	$V_{dc}/2$
32	1	1	1	0	1	1	$V_{dc}/3$	$V_{dc}/4$
33	1	1	1	1	1	1	0	0

Sources: Authors Compilation.

Figure-4: Adjacent Switching Vectors for a Seven-Level Diode-Clamped Inverter with Symmetrical and Asymmetrical DC Link Configurations.



Sources: Authors Compilation.

From the possible switching function states defined in Table I, Figure4 shows the adjacency diagram of SDCC and ADCC config-urations. As shown, the adjacent vectors are available between all voltage levels in both configurations, so that all voltage levels can be achieved with one switch change. However, nonadjacent switching transitions are required between the following switch-ing transitions in the ADCC configuration, which are depicted by dashed lines in Fig. 4:

1) (20) and (31)
2) (02) and (13).

Table-3: Maximum Voltage Rate across Switching Components in One Leg of a Seven-Level H-Bridge Diode-Clamped Inverter with Symmetrical and Asymmetrical DC Link Arrangements

Switching components	Max. voltage rating of SDCC	Max. voltage rating of ADCC
S_1	$V_{dc}/3$	$V_{dc}/3$
S_2	$V_{dc}/3$	$V_{dc}/3$
S_3	$V_{dc}/3$	$V_{dc}/2$
S_4	$V_{dc}/3$	$V_{dc}/3$
S_5	$V_{dc}/3$	$V_{dc}/3$
S_6	$V_{dc}/3$	$V_{dc}/2$
D_{c1}	$V_{dc}/3$	$V_{dc}/4$
D_{c2}	$2V_{dc}/3$	$3V_{dc}/4$
D_{c3}	$2V_{dc}/3$	$V_{dc}/2$
D_{c4}	$V_{dc}/3$	$V_{dc}/2$

Sources: Authors Compilation.

This transition requires two switch changes in the ADCC. To remove nonadjacent voltage vector in positive voltage levels, when the controller increases the voltage level from $V_{dc}/2$ to $3V_{dc}/4$, transition occurs from switching function states (31) to (20). Then, the controller uses switching function state (10) for modulation between $V_{dc}/2$ and $3V_{dc}/4$. Also, after the oc-currence of the transition from $3V_{dc}/4$ (20) to $V_{dc}/2$ (31), the

controller uses state (21) for the modulation between $V_{dc}/2$ and $V_{dc}/4$. The same situation happens when the output voltage is negative.

Therefore, these nonadjacent switching transitions occur only four times during one cycle. It is apparent that by assuming that switching losses are proportional to the number of switching per cycle, the switching loss associated with the extra switching is negligible at high switching frequency. As a re-sult, improved voltage waveforms can be obtained using ADCC topology compared with SDCC topology with the same number of components and almost the same number of switchings. Although the output voltage of the asymmetrical dc link ar-rangement benefits from two more voltage levels with the same number of components compared with the symmetrical dc link arrangement, extra voltage rates should be taken into account for two switches in each leg of the inverter. Maximum voltage across switching components during different switching states is derived in TableIII to have a comparison between the volt-age rating in the symmetrical and asymmetrical dc link arrange-ments.

In the asymmetrical configuration, the maximum voltage rating of switches (S_3 and S_6) in each leg is $V_{dc}/6$ more than the switches in the symmetrical configuration for the same dc link voltage. By investigating the voltage ratings, maximum voltage rating of diodes (D_{c1} and $D_{c\ 3}$) decreased by $V_{dc}/12$ and $V_{dc}/6$; however, the maximum voltage tolerated by another two diodes (D_{c2} and $D_{c\ 4}$) increased by $V_{dc}/12$ and $V_{dc}/6$, which shows that the maximum voltage across diodes has not changed in both configurations.

The output voltage of the symmetrical and asymmetrical single-phase diode-clamped inverters for the same circuit pa-rameters is shown in Fig. 5. Herein, at the dc–dc side, input volt-age (E) is assumed to be 100 V; switching frequency of the dc–dc converter (f_{sw}) is 10 kHz, L = 2 mH, and $C_1 = C_2 = C_3$ = 1 mF, while at the inverter side, fundamental and switching fre-quencies are f = 50 Hz, f_{sw} = 4 kHz, and the dc link of the seven-level diode-clamped inverter (V_{NM}) is boosted to 300 V using a triple-output boost converter. Midpoint voltage regulation for symmetrical and asymmetrical configurations for high modulation index (m_a = 1) and unity power factor (PF) has been shown in Fig. 5(a).

It is clear that the MOB converter is able to boost the low input voltage for dc link capacitors as well as balance the capacitors voltage to the desired level for m_a = 1 and pure resistive load, which is impossible in more than five-level single-phase diode-clamped topology without an active front-end converter. In order to generate the output volt-age, based on the duty cycle of switches, the controller chooses the next suitable switching function state using the adjacent vectors in Fig. 4.

To show the performance of the proposed structure for inductive load, Fig. 5(b) illustrates the dc link capacitor voltage control and output voltage for m_a = 1 and PF = 0.5. In order to synthesize an equal dc link capacitor voltage arrangement in the conventional configuration, while the total voltage of an inverter dc link is boosted at 300 V, midpoint voltages (V_{c1} and $V_{c1} + V_{c2}$) are controlled at 100 and 200 V.

However, to have an asymmetrical dc link config-uration, midpoint voltages are controlled at 150 and 225 V, respectively.

A laboratory prototype of a symmetrical and an asymmetrical seven-level H-bridge inverter has been implemented to practi-cally verify the proposed configuration.

The laboratory proto-type has been tested for the following specifications: V_{dc} = 90 V, $I_{ut\text{-}peak}$ = 5 A, f = 160 Hz, and f_{sw} = 6 kHz under pure in-ductive load L = 16 mH. A predictive current control has been developed in a V850E/IG3 microcontroller to force the load cur-rent to follow the for the H-bridge seven-level diode-clamped inverter with symmetrical and asymmetrical dc link arrange-ments.

Switching states to generate the desired voltage level based on the amount of current reference is chosen by the mi-crocontroller according to adjacent switching vectors. Output voltage and current of the symmetrical and the asymmetrical seven-level diode-clamped inverter is demonstrated in Fig. 5.

Regarding simulation and hardware results, two more voltage levels can be synthesized in output voltage of the ADCC config-uration compared with the SDCC configuration with the same number of components and structure. Therefore, the seven-level H-bridge inverter performs similarly to a nine-level inverter.

To examine the performance of the asymmetrical configuration, harmonic spectrums associated with the output voltage of both the SDCC and ADCC configurations are exposed. Comparing harmonic spectrums, better harmonic performance is obtainable using the asymmetrical dc link arrangement for the diode-clamped converter.

This achievement allows, on one hand, an improvement in output voltage harmonic characteristics with the same number of components compared with the symmetri-cal seven-level H-bridge configuration, and on the other hand, a decrease in cost and complexity of the inverter hardware lay-out structure with the same output waveforms quality compared with the symmetrical nine-level H-bridge inverter.

III. PROPOSED MULTILEVEL HYBRID CASCADE CONVERTERS

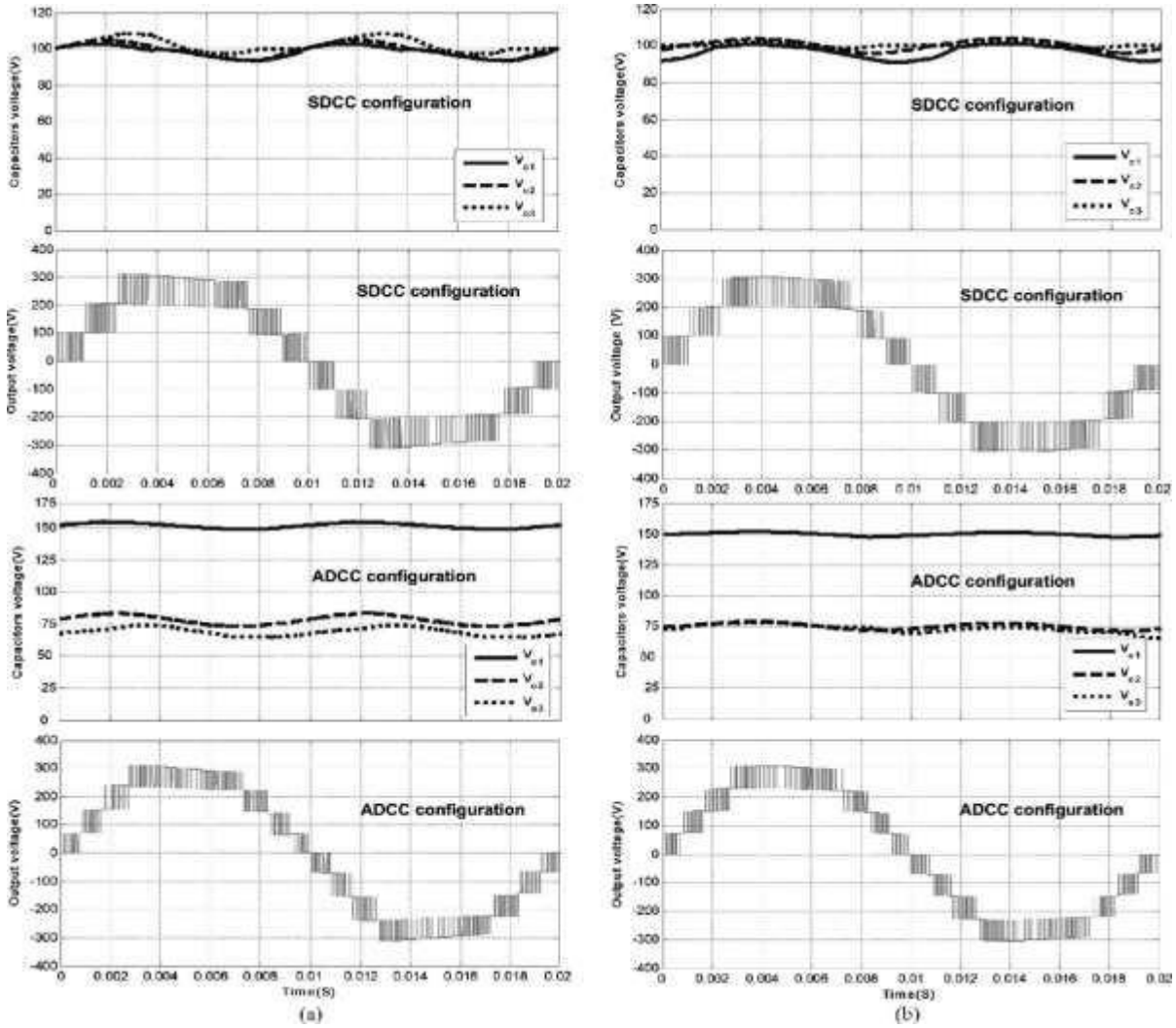
One of the aspects of this paper is to select dc input voltages for a topology based on a series connection of a symmetrical and an asymmetrical diode-clamped H-bridge cell with three-level H-bridge inverters to achieve a maximum number of output lev-els by preserving the minimum switching losses.

Simultaneous switching of different switches is not a real problem when there are just few of them happening over one cycle; however, repeat-edly switching between the nonadjacent switching vectors is not acceptable, due to an increase in switching losses and com-mutation noise.

The state of art in this topology can improve the resolution of the output voltage with minimum power com-ponents, while keeping the adjacent switching vectors between all modulation voltage levels.

Fig. 4 presents the schematic of this configuration for a two-cell hybrid cascade converter. The MOB converter is utilized to supply dc input voltage of the diode-clamped multilevel H-bridge cell to regulate capacitors voltages and provide the desired voltage rate for the dc link capacitors.

Figure-5: DC link capacitor voltages and output voltage for SDCC configuration and ADCC configuration with (a) resistive load (m_a = 1, PF = 1) and (b) inductive load (m_a = 1, PF = 0.5)



Source: NAMI *et al.*: Hybrid Cascade Converter Topology.

To achieve maximum resolution in output voltage of the N-cell proposed topology, the dc voltage arrangement should be considered as follows:

$$V_{dcN} = \frac{\text{Minimum voltage level of multilevel inverter}}{2^{N-1}} \quad (1)$$

where, for the proposed two-cell inverter in Fig. 7, we have

$$V_{dc2} = \frac{\text{Minimum voltage level of multilevel inverter}}{2} \quad (2)$$

Based on this arrangement, the number of output voltage levels can be derived from (1)

$$\text{number of output levels} = 2^{N-1} \times (m+1) - 1 \quad (3)$$

where m is the number of output voltage levels of multilevel H-bridge inverters and N is the number of cells. Regarding (3), there are two possibilities to increase the obtainable voltage lev-els in the proposed configuration.

One solution is by adding a classical three-level H-bridge inverter in which the number of voltage levels will be doubled. This imposes four extra switches.

IV. CONCLUSIONS

This paper has presented the diode-clamped multilevel Hbridge cell cascaded with three-level conventional inverters to increase efficiency of converters with high output voltage resolution. A novel dc link voltage rating is proposed for the multilevel diode-clamped and three-level H-bridge inverters to improve the output voltage and current quality by preserving the adjacent switching vectors between all voltage levels. The MOB converter has been applied as a dc link supplier of a diode clamped inverter to boost and regulate the capacitors' voltage to the desired dc link rates. Using the MOB converter, a new cascade inverter is verified by cascading asymmetrical seven-level H-bridge diode-clamped inverter. Inverter with nineteen output voltage levels performance was achieved, which has more voltage levels as well as lower voltage, and current THD rather than using a symmetrical diode-clamped inverter with the same configuration and equivalent number of power components. Predictive current control was conducted to

show the performance of the proposed method. In this case, two different methods for the switching states selection are proposed to minimize either losses or THD of voltage in hybrid converters. Novel H-bridge cascaded cells will decrease the complexity of control and cost of the system as well as diminish or remove the output filters when the configuration will be extended for more H-bridge cells.

REFERENCES

[1] A., Nabae; I., Takahashi, and H., Akagi, (1981). "A New Neutral-point-clamped PWM Inverter", IEEE Trans. Ind. Appl., September / October, Volume: 1A-17, Number; 5, pp. 518–523.

[2] M., F. Escalante; J., C. Vannier, and A., Arzande, (2002). "Flying Capacitor Multilevel Inverters and DTC Motor Drive Applications", IEEE Trans. Ind. Electron., August, Volume: 49, Number: 4, pp. 809–815.

[3] F., Z. Peng; J., W. McKeever, and D., J. Adams, (177). "Cascade Multilevel Inverters for Utility Applications", Proceeding IEEE IECON, Volume: 2, pp. 437–442.

[4] L., G. Franquelo; J., Rodriguez; J., I. Leon; S., Kouro; R., Portillo, and M., A. M. Prats, (2008). "The Age of Multilevel Converters Arrives", IEEE Ind. Electron. Mag., June, Volume: 2, Number: 2, pp. 28–39.

[5] S., B. Kjaer; J., K. Pedersen, and F., Blaabjerg, (2005). "A Review of Single-Phase Grid-Connected Inverters for Photovoltaic Modules", IEEE Trans. Ind. Appl., September / October, Volume: 41, Number: 5, pp. 4354–4361.

[6] J., M. Carrasco; L., G. Franquelo; J., T. Bialasiewicz; E., Galv´an; R., C. P. Guisado; Ma., A. M. Prats; J., I. Leo´n, and N., Moreno-Alfonso, (2006). "Powerelectronic Systems For The Grid Integration of Renewable Energy Sources: A Survey", IEEE Trans. Ind. Electron., August, Volume: 53, Number: 4, pp. 1002–1016.

[7] Zh., Du; L., M. Tolbert; J., N. Chiasson, and B., Ozpineci, (2006). "A Cascade Multilevel Inverter Using a Single DC Source", Proceedings 21st Annual IEEE APEC, March, pp. 426–430.

[8] S., Vazquez; J., I. Leon; L., G. Franquelo; J., J. Padilla, and J., M. Carrasco, (2009). "DC-Voltage-Ratio Control Strategy for Multilevel Cascaded Converters Fed with a Single DC Source", IEEE Trans. Ind. Electron., July, Volume: 56, Number: 7, pp. 2513–2521.

[9] M., Veenstra, and A., Rufer, (2005). "Control of a Hybrid Asymmetric Multilevel Inverter for Competitive Medium-Voltage Industrial Drives", IEEE Trans. Ind. Appl., March / April, Volume: 41, Number: 2, pp. 655–664.

[10] E., Babaei, (2008). "A Cascade Multilevel Converter Topology with Reduced Number of Switches", IEEE Trans. Power Electron, November, Volume: 23, Number: 6, pp. 2657–2664.

[11] A., Nami; F., Zare; G., Ledwich; A., Ghosh, and F., Blaabjerg, (2008). "Comparison between Symmetrical and Asymmetrical Single Phase Multilevel Inverter with Diode-Clamped Topology," Proceedings IEEE PESC 2008, pp. 2921–2926.

[12] J., Song-Manguelle, (2001). "Asymmetrical Multilevel Inverter for Large Induction Machine Drive", Electrical Drives and Power Electronics International Conference, The High, Tatras, Slovakia.

[13] J., D. L. Morales; M., F. Escalante, and M., T. Mata-Jimenez, (2007). "Observer for DC Voltages in a Cascaded H-Bridge Multilevel STATCOM", IET Electr. Power Appl., Volume: 1, pp. 879–889.

[14] R., Naderi, and A., Rahmati, "Phase-shifted carrier PWM Technique for General Cascaded Inverters", IEEE Trans. Power Electron., May, Volume: 23, Number: 3, pp. 1257–1268.

[15] J., Rodriguez; L., Moran; P., Correa, and C., Silva, (2002). "A Vector Control Technique for Medium-Voltage Multilevel Inverters", IEEE Trans. Power Electron, August, Volume: 49, Number: 4, pp. 882–888.

[16] M., S. A. Dahidah, and V., G. Agelidis, (2008). "Selective Harmonic Elimination PWM Control for Cascaded Multilevel Voltage Source Converters: A Generalized Formula", IEEE Trans. Power Electron, July, Volume: 23, Number: 4, pp. 518–529.

[17] J., Rodr´ıguez; S., Bernet; B., Wu; J., O. Pontt, and S., Kouro, (2007). "Multilevel Voltage-Source-Converter Topologies for Industrial Medium-Voltage Drives", IEEE Trans. Ind. Electron, December, Volume: 54, Number: 6, pp. 2930–2945.

[18] C., Rech, and J., R. Pinheiro, (2007). "Impact of hybrid Multilevel Modulation Strategies on Input and Output Harmonic Performances", IEEE Trans. Power Electron., May, Volume: 22, Number: 3, pp. 967–977.

[19] Zh., Du; L., M. Tolbert; B., Ozpineci, and J., N. Chiasson, (2009). "Fundamental Frequency Switching Strategies of a Seven-Level Hybrid Cascaded H-Bridge Multilevel Inverter", IEEE Trans. Power Electron., January, Volume: 24, Number: 1, pp. 25– 33.

[20] M., D. Manjrekar; P., K. Steimer, and T., A. Lipo, (2000). "Hybrid Multilevel Power Conversion System: A Competitive Solution for High-Power Applications", IEEE Trans. Ind. Appl., May / June, Volume: 36, Number: 3, pp. 834–841, 2000.

[21] N., S. Choi; J., G. Cho, and G., H. Cho, (1991). "A General Circuit Topology of Multilevel Inverter", Proceedings IEEE PESC' 1991, pp. 96–103.

[22] C.,-C. Hua; C.,-W. Wu, and C.,-W. Chuang, (2006). "Control of Low-Distortion 27-Level Cascade Inverter with Three H-Bridge Inverter Modules", Proceedings IEEE ICIT, pp. 277–282.

[23] M., N. Abdul Kadir, and Z., F. Hussien, (2004). "Asymmetrical Multilevel Inverter Using Ratio-Three Based Sources", Proceedings IEEE PECon pp. 137–142.

*****

CONDITION MONITORING ON VIBRATION RESPONSE OF MISALIGNED ROTORS

Jeevanantham P.[91] Jayaprakash S.[92]

ABSTRACT

The accuracy of modeling of rotor systems composed of rotors, bearings and a foundation is evaluated and discussed in this paper. The experimental data have been obtained with a fully instrumented three ball bearing, two shafts test rig. The fault models are then used in the frame of a model based malfunction identification procedure in the time domain and frequency domain. Rotating machines are extensively used in diverse engineering applications, such as power stations, marine propulsion systems, aircraft engines, machine tools, automobiles and other accessories. The vibration in rotating machinery is mostly caused by unbalance, misalignment, mechanical looseness, shaft and other malfunctions. All these malfunctions can lead to a premature failure of a bearing, the shaft or coupling. Condition monitoring is a valuable tool to extend the operating life of a machine. The various conditions monitoring techniques available vibration monitoring is the most widely used one in industry today. The objective of vibration monitoring is the detection of changes in the vibration condition of the object under investigation during its operation by using lab view. The cause of such changes is mainly the appearance of any of the malfunctions.

KEYWORDS

Unbalance, Misalignment, Detection, Vibration, Mechanical Looseness etc.

I. INTRODUCTION

The most common cause of machine vibration is rotor misalignment after imbalance, which leads additional dynamic load and accelerate machine deterioration. This type of vibration is often from reactive forces in the couplings between two rotating shafts. Understanding and practicing the fundamentals of rotating shaft parameters is the first step in reducing this unnecessary vibration, reducing maintenance costs and increasing machine uptime. In industry, 30% of a machine's downtime is due to the poorly aligned machines [1]. Misalignment is estimated to cause over 70% of rotating machinery's vibration problems [3]. Hence, an in-depth study and an accurate knowledge on the vibration characteristics is very helpful in understanding and diagnosing the rotor misalignment to avoid any failures or damages that may arise [2]. There are various factors such as differential thermal growth of machines, asymmetry in applied loads, unequal foundation settlement, etc., which disturb the alignment condition of a machine [2].

Despite the rapid increase in the understanding of rotor dynamics, no satisfactory analysis explains the range of observed phenomena [1]. Vance [4] and Goodman [5] observed that misalignment is present due to improper machine assembly and sometimes, the thermal distortion of the bearing housing supports, resulting in abnormal rotating preload. However, a perfect alignment between the driving and driven shafts cannot be attained. Gibbons [6] first derived the misalignment reaction forces from those generated in different types of couplings

Xu and Marangoni [7] showed analytically that the vibration due to coupling misalignment mainly occur at the even multiples of the rotor speed. Sekhar and Prabhu [8] numerically evaluated the effect of coupling misalignment on the vibration response of the rotor. They suggested 2X vibration response as a characteristic signature of misaligned shafts. Dewell and Mitchell showed experimentally that 2X and 4X vibration components are largely dependent upon coupling misalignment. Lee [9] derived a model for the flexible coupling–rotor–ball bearing system, including reaction loads from deformations of rolling elements of bearing and coupling elements as the misalignment effects. From orbital analysis, anisotropy of bearing stiffness was suggested as the misalignment indicator. The other approach often used to simulate misalignment effect in rotor system is from kinematics of the couplings. Piotrowski concluded that vibration due to misalignment is usually characterized by a 2X running speed component and high axial vibration levels. When a misaligned shaft is supported by rolling element bearing, these characteristic frequencies may also appear. Tejas and Ashish [2] found from the measured forces that the presence and type of misalignment (parallel and angular misalignment) has significant influences on the harmonic content of the misalignment excitation forces.

From the literature, it is clearly proven that misalignment produces high vibration levels in bearings. It is influenced by the machine speed and the stiffness of the coupling. e.g., rubber couplings are more tolerant and tend to produce less amount of vibration. In general, the majority of misalignment studies in the past are theoretical whereas experimental investigations are relatively limited. The theoretical studies are often investigated with simplified assumptions. The outcome of these studies may not be accurate, since in practice there are many more sources of observed vibration characteristics in an actual rotor system. Moreover, all of these studies used wired accelerometer which is usually attached to bearings housing. In addition, it focuses specifically on experimental study of misaligned coupled rotors supported by ball and roller bearings.

$$\text{Torque}=60P/2\Pi N \quad (1)$$

$$\text{Moment}=wl/4 \quad (2)$$

$$\text{Combined Moment } (M_e)=1/2[(2xm)+V[(2xm)^2+(1.5xt)^2]] \quad (3)$$

$$\text{Stiffness } (k) =48EI/L^3 \quad (4)$$

$$\text{Natural frequency}=V\,(k/m) \quad (5)$$

$$\text{Angular frequency } (w) =2\Pi N/60 \quad (6)$$

II. TESTING RING

The experimental setup consists of a motor of 1/12 hp. Two shafts of diameter 12mm are used supported by three pillow block bearings. Two types of couplings are employed, love joy coupling and self design coupling. Lovejoy coupling acts as vibration damper and isolator. Self design coupling is used with no fault and three different parallel misalignments of 1mm,

[91] Assistant Professor, Kalaignar Karunanidi Institute of Technology, Tamil Nadu, India, jeevananthamed@gmail.com
[92] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, jpmecadcam@gmail.com

1.5mm and 2mm. balanced mass condition and unbalanced mass condition are analyzed. Vibration signals are acquired through accelerometer which is placed on top of the three different bearings. From these signals displacement, velocity and acceleration are analysed.RMS and peak to peak values are taken for different misalignments and different speeds. The tachometer is used to measure the speed. Pillow types of bearings such asNBUP204 UC201KBL are used.

Figure-1: Test Ring

Sources: Authors Compilation.

Table-1

Parameter	Factor
1.First coupling	Zero offset
2.second coupling	1mm offset
3.Third coupling	1.5mm offset
4.Fourth coupling	2mm offset
5.Balanced mass	870grams
6.Balancedmass+ Un balanced mass	870gram+40grams

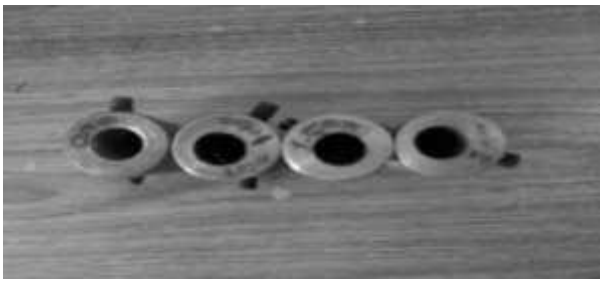
Sources: Authors Compilation.

III. EXPERIMENTAL DETAILS

To objective of this experiment to analysis the vibration signature of misaligned rotor and unbalanced rotor at various conditions like 1mm, 1.5mm and 2mm offset Analysis performed in different speed condition like 600, 700, 800, 900, 1000 rpm with mass added 870grams and unbalanced 40 grams condition.

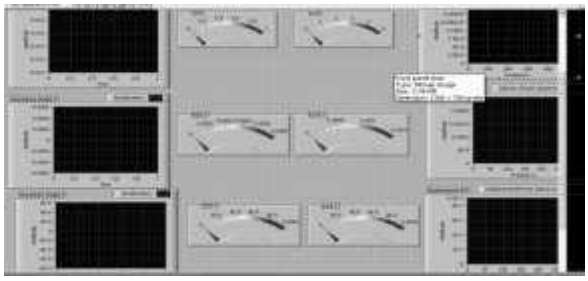
A typical condition monitoring system has been developed using LABVIEW software with sound and vibration tool kit. An effective continues signal extraction system is developed for monitoring of misaligned rotor and un balanced rotor signatures. The extracted signals are diagnosed with time domain analysis like RMS and PEAK to PEAK and frequency domain analysis like: power spectrum analysis.

Figure-2: 0, 1, 1.5, 2 MM Offset Couplings



Sources: Authors Compilation.

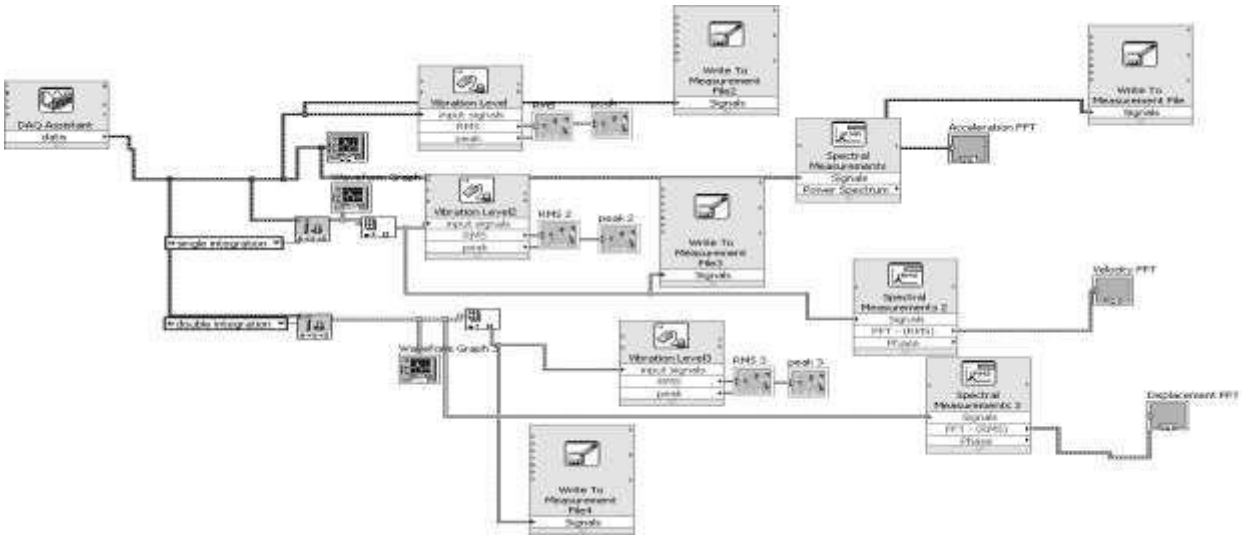
Figure-3: Lab View Front Panel



Sources: Authors Compilation.

Initially zero offset coupling was fixed in the test rig and signals were recorded using DAQ (data Acquisition). The zero offset coupling was replaced by 1, 1.5 and 2 mm offset couplings and signals were recorded for each one of the case separately under the same condition. Time domain analysis and frequency domain were performed in both signals. Here the figure of front panel and block diagram are shown below:

Figure-4: Lab View Block Diagram

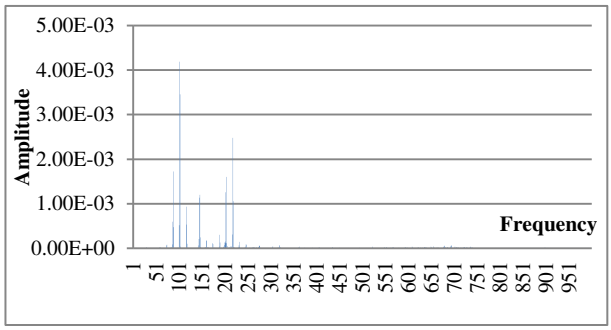


Sources: Authors Compilation.

IV. PREMATURE OF TEST RIG VIBRATION

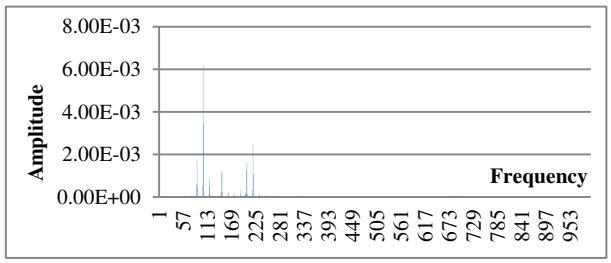
Prior to vibration analysis 3 different bearings are examined on its vibration levels.

Figure-5: Nofault Condition First Bearing at 600rpm



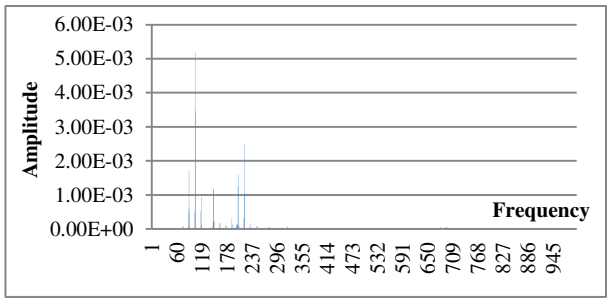
Sources: Authors Compilation.

Figure-6: Nofault Condition Second Bearing at 600rpm



Sources: Authors Compilation.

Figure-7: No-Fault Condition Third Bearing at 600rpm



Sources: Authors Compilation.

By comparing three bearing vibration signals, it is clear that the second bearing was the highest vibration level because of self designed coupling gives more vibration. First bearing produces the less vibration compared to second bearing; first bearing nearby using love joy coupling. Third bearing also less vibration compared to second bearing vibration level.

V. RESULTS AND DISCUSSIONS

Following conclusions are arrived from the experiment. The comparison of vibration signals of zero offset and 1, 1.5 and 2mm offset of couplings respectively i.e. balanced mass adding condition and unbalanced mass adding condition. The displacement, velocity, and acceleration of RMSandPEAK predicted. The comparison of vibration signals in acceleration RMS and PEAK to PEAK value to be analyzed. Zero offset of coupling to be taken no fault condition. 1, 1.5, and 2mm parallel misalignment to be taken fault condition. Compared to balance mass condition and un-balanced mass condition to be analyzed. Here graph shows the acceleration of RMS value to be taken to no fault condition reference value. Fault condition of test rig induces misalignment of coupling near second bearing of RMS value should be high in misaligned condition. Unbalanced condition of RMS value should be very high compared to balanced mass condition of RMS value. Graph shows misalignment and unbalance increased means RMS value also increased.

Table-2: No-Fault

(1)No fault condition first bearing RMS value			
n[rpm]	z[mm]	z[mm/s]	z[mm/s^2]
600	3.00E-05	0.00025	0.1
700	3.00E-05	0.00025	0.125
800	3.50E-05	0.000275	0.15
900	3.50E-04	0.00032	0.2
1000	5.00E-04	0.0007	0.4
(2)No fault condition second bearing RMS value			
600	8.00E-04	0.00035	0.135
700	9.50E-04	0.0004	0.155
800	2.00E-03	0.00055	0.2
900	3.00E-03	0.0006	0.4
1000	3.50E-03	0.0009	0.65
(3)NO fault condition third bearing RMS value			
600	3.00E-05	0.0003	0.125
700	3.50E-05	0.00035	0.135
800	4.00E-05	0.0004	0.18
900	5.00E-04	0.00045	0.3
1000	2.00E-03	0.0008	0.5

Sources: Authors Compilation.

Table-3: 1mm Offset

(1)1mm parallel misalignment first bearing RMS value			
n[rpm]	z[mm]	z[mm/s]	z[mm/s^2]
600	3.00E-04	0.00025	0.2
700	3.50E-04	0.00025	0.25
800	0.0005	0.00035	0.35
900	7.00E-04	0.035	0.7
1000	0.004	0.04	1
(2)1mm parallel misalignment second bearing RMS value			
600	3.50E-03	0.04	0.3
700	4.00E-03	0.045	0.35
800	0.005	0.055	1.25
900	0.006	0.06	1.5
1000	0.007	0.07	1.75
(3)1mm parallel misalignment third bearing RMS value			
600	2.50E-03	0.028	0.225
700	3.50E-03	0.032	0.275
800	4.00E-03	0.038	0.9
900	4.50E-03	0.043	1
1000	5.00E-03	0.052	1.25

Sources: Authors Compilation.

Table-4: 1.5 mm Offset

(1)1.5mm parallel misalignment first bearing RMS value			
n[rpm]	z[mm]	z[mm/s]	z[mm/s^2]
600	3.00E-04	0.0007	0.25
700	3.50E-04	0.0009	0.3
800	0.002	1.00E-02	0.8
900	0.003	0.02	0.825
1000	0.007	0.05	1.25
(2)1.5mm parallel misalignment second bearing RMS value			
600	3.00E-03	0.001	0.45
700	4.00E-03	0.00125	0.55
800	5.00E-03	0.004	1.35
900	0.006	0.0045	1.75
1000	0.009	0.07	2.75
(3)1.5mm parallel misalignment third bearing RMS value			
600	1.00E-03	0.0009	0.3
700	3.00E-03	0.00095	0.35
800	5.00E-03	0.001	1
900	0.006	0.002	1.25
1000	0.008	0.003	1.35

Sources: Authors Compilation.

Table-5: 2 mm Offset

(1)2mm parallel misalignment first bearing RMS value			
n[rpm]	z[mm]	z[mm/s]	z[mm/s^2]
600	3.00E-03	0.001	0.275
700	4.20E-03	0.005	0.35
800	6.50E-03	0.008	0.9
900	0.007	0.05	1.25
1000	0.009	0.06	1.5
(2)2mm parallel misalignment second bearing RMS value			
600	2.00E-02	0.002	0.7
700	3.00E-02	0.0035	0.75
800	5.00E-02	0.0045	1.5
900	8.00E-02	0.0055	2.5
1000	9.50E-02	0.09	3.5
(3)2mm parallel misalignment third bearing RMS value			
600	7.00E-03	0.0015	0.5
700	7.50E-03	0.006	0.55
800	0.008	0.007	1.25
900	0.0085	0.07	2
1000	0.0009	0.09	2.5

Sources: Authors Compilation.

Table-6: Balanced

(1)Balanced condition first bearing RMS value			
n[rpm]	z[mm]	z[mm/s]	z[mm/s^2]
600	3.50E-03	0.00025	0.125
700	4.00E-03	0.00028	0.14
800	4.50E-03	0.00032	0.275
900	5.00E-03	0.00045	0.325
1000	6.00E-03	0.00056	1.25
(2)Balanced condition second bearing RMS value			
600	4.50E-03	0.000325	0.14
700	5.50E-03	0.000375	0.275
800	6.20E-03	0.000475	0.34
900	7.50E-03	0.00055	0.4
1000	8.00E-03	0.0007	1.4
(3)Balanced condition third bearing RMS value			
600	4.00E-03	0.0003	0.16
700	4.75E-03	0.00035	0.25
800	5.50E-03	0.00045	0.32
900	6.00E-03	0.0005	0.38
1000	6.50E-03	0.00055	1.3

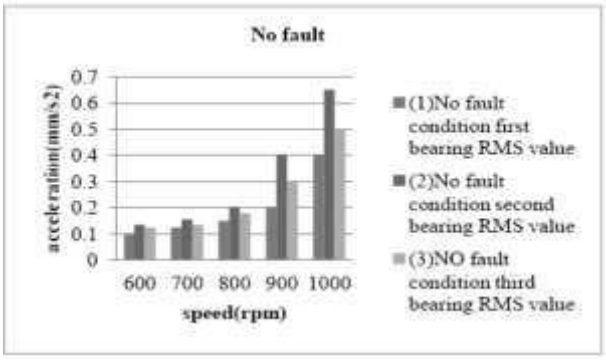
Sources: Authors Compilation

Table-7: Un-Balanced

(1)unBalanced condition first bearing RMS value			
n[rpm]	z[mm]	z[mm/s]	z[mm/s^2]
600	3.75E-02	0.00035	0.14
700	4.00E-02	0.00042	0.275
800	4.50E-02	0.00048	0.3
900	5.50E-02	0.00052	0.375
1000	5.75E-02	0.00065	1.5
(2)un Balanced condition second bearing RMS value			
600	5.00E-02	0.000045	0.18
700	6.20E-02	0.0009	0.325
800	6.90E-02	0.001	0.57
900	7.20E-02	0.003	0.625
1000	9.20E-02	0.004	1.9
(3)un Balanced condition third bearing RMS value			
600	4.50E-02	0.0004	0.17
700	5.00E-02	0.0008	0.3
800	5.50E-02	0.0009	0.5
900	6.20E-02	0.001	0.52
1000	7.20E-02	0.002	1.95

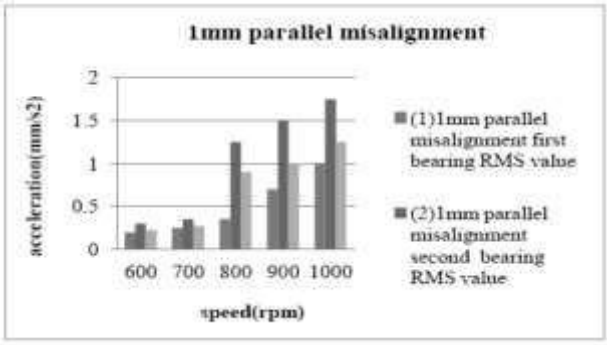
Sources: Authors Compilation

Graph-1



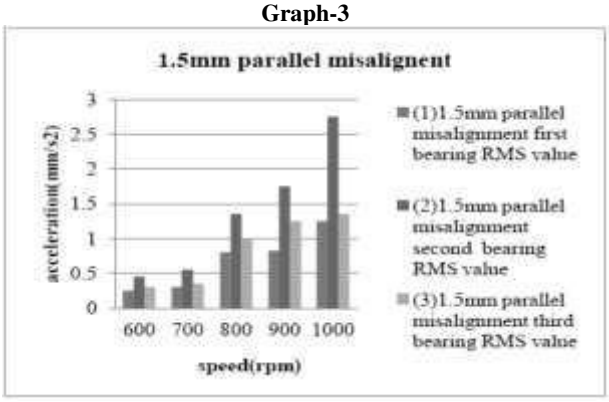
Sources: Authors Compilation.

Graph-2



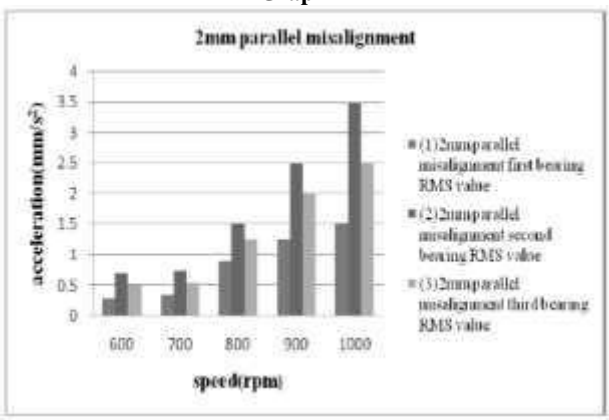
Sources: Authors Compilation.

Graph-3



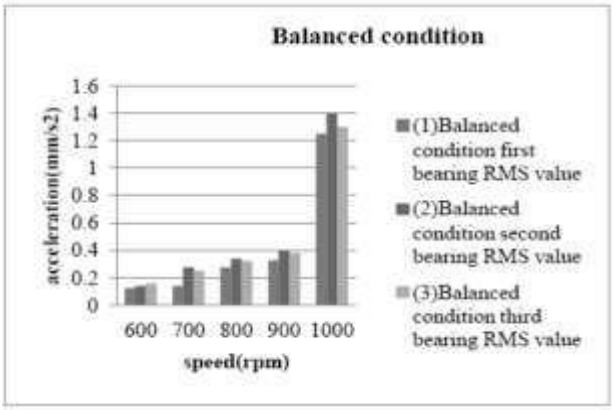
Sources: Authors Compilation.

Graph-4



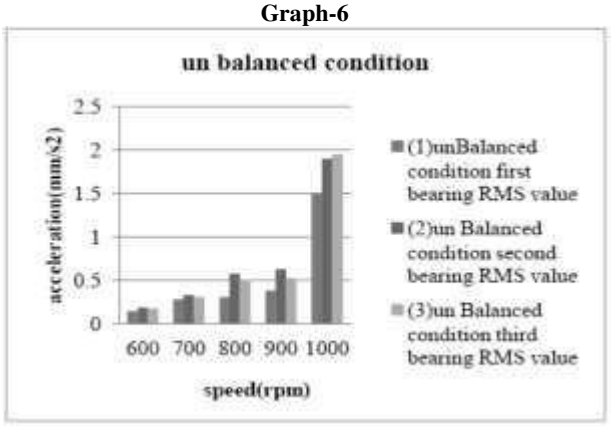
Sources: Authors Compilation.

Graph-5



Sources: Authors Compilation.

Graph-6



Sources: Authors Compilation.

CONCLUSIONS

The experimental results demonstrated that the vibration monitoring rig modelled various modes of machine failure is indeed capable of independently generate common machine faults. With this test rig common machine faults can be introduced and the results are analysed and thus life of the component is predicted. Self-designed couplings were used in the experiments. The shaft alignments were performed on the test ring. The frequency spectra were obtained for the results.

The theoretical predictions are in good agreement with the experimental measurements. The experimental arrangement shows that the different couplings exhibit different frequency characteristics in the vibration spectra. This phenomenon can be used in predictive maintenance to monitor the shaft conditions.

Also, the vibration spectrum can be analysed to determine the machine conditions such as misalignment and unbalance. Then repair or maintenance can be scheduled if necessary to avoid unexpected machine failures and costly machinery downtime.

REFERENCES

[1] Tejas, H. Patel, and Ashish K. Darpe, (2009), "Experimental Investigations on Vibration Response of Misaligned Rotors", Mechanical Systems and Signal Processing, Volume; 323, pp.2236–2252.

[2] Tejas, H. Patel, and Ashish K. Darpe, (2009), "Vibration Response of Misaligned Rotors", Journal of Sound and Vibration, Volume: 325, pp. 609–628.

[3] Arun, Kr. Jalan, and A., R. Mohanty, (2009), "Model Based Fault Diagnosis of a Rotor–Bearing System for Misalignment and Unbalance under Steady-State Condition", Journal of Sound and Vibration, Volume: 325, pp. 604-622.

[4] Surendra, N.; Mark, H., and Zhuang Li, (2008), "Using Operating Deflection Shapesto Detect Misalignment in Rotating Equipment", Presented at IMAC XXVI.

[5] A., W. Lees; J., K. Sinha, and M., I. Friswell, (2003), "Estimating Rotor Unbalance and Misalignment from a Single Run-Down", 5th International Conference on Modern Practice in Stress and Vibration.

[6] S., Generiwala; Patel, H., and A., Hartung, (2007), "Observation Concerning Misalignment Vibration Signatures", Spectra Questinc.

[7] M., Xut, and R., Dmarangoni, (1993), "Vibration Analysis of A Motor – Flexible Coupling – Rotor System Subject To Misalignment and Unbalance Part1: Theroetical Model and Analysis", Journal of Sound and Vibration, Volume: 76, pp. 663-679.

[8] M., Xut, and R. Dmarangoni, (1993), "Vibration Analysis of A Motor – Flexible Coupling – Rotor System Subject To Misalignment and Unbalance Part2: Experimental Validation", Journal of Sound and Vibration, Volume: 76, pp. 663-679.

[9] Menderes, Kalkat; Sahin, Yildirim, and Ibrahim, Uzmay, (2004), "Design of Artificial Neural Networksfor Rotor Dynamics Analysis of Rotating Machine System", Journal of the Brazilian Society of Mechanical Sciences, Volume: 15, pp. 573-588.

[10] Nicolò, Bachschmid; Paolo, Pennacchi; Ezio, Tanzi, and Andrea, Vania, (2000), "Accuracy of Modeling and Identification of Malfunctions in Rotor Systems: Experimental Results", Journal of Brazilian Society of Mechanical Sciences, Volume: 22, Number: 3.

*****

SYNTHESIS AND MAGNETIC PROPERTIES OF NICKEL NANO PARTICLES

M. Hemalatha[93] N. Suryanarayanan[94] S.Prabahar[95]

ABSTRACT

Nickel nano particles around 12nm were prepared by chemical reduction of Nickel Chloride dissolved in Ethylene glycol, in which hydrazine act as a reducing agent and NaOH as stabilizer. The powder XRD study confirms that the average size of the particle is 12nm.The Ni particles show the characteristic of super paramagnet with saturation magnetization (M_s) and remanent magnetization (M_r) at 5K is 2.55 emu/g and 1.11 emu/g, respectively.

KEYWORDS

Synthesis, Nano Particles, Nickel, Chemical Reduction Method etc.

I. INTRODUCTION

In the past two decades, considerable attention has been devoted to the synthesis of metal nano particles because of their unusual properties and potential applications in optical, electronic, catalytic and magnetic materials [1-3]. Until now, only a few works on the preparation of Nickel nano particles have been reported and they usually were performed in Organic media to avoid the formation of nickel oxide or hydroxide [4].

Szu-Han Wu et al [5] and Yang long Hou et al [6] prepared Nickel nanoparticles by chemical reduction method and characterized Nickel to be pure crystalline with a FCC structure. Xiao-Min Ni et al [7] had prepared the Nickel nano rods in Water-in-oil micro emulsion technique with a FCC structure. Kumar et al [10] investigated that Nickel nanoparticles show a typical behaviour for their field-dependent magnetization.

Kumar et al [8] investigated that Nickel nanoparticles show a typical behaviour for their field-dependent magnetization. Yingwen Duan et al [9] have synthesized Ni nanoparticles by spontaneous autocatalytic reduction and found that the specific saturation magnetization decreases as the grain size decreases.Our objective in this paper is to synthesize nickel nanoparticles by chemical reduction method using nickel chloride and hydrazine in an aqueous solution.

II. EXPERIMENTAL DETAILS

Ethylene glycol was chosen as a solvent and $NiCl_2$ was dissolved in it. Then, an appropriate amount of hydrazine was added. Hydrazine act as a reducing agent and NaOH, which acts as a catalyst of 10-72 μ/ml, were added in sequence and stirred well.

The resultant product was taken and kept in the water bath above the room temperature and heated. No significant change takes place. When the temperature raised to above 90° C, there was a change in the resultant product. That is, the bluish green color precipitate slowly changed in to black color, indicating the formation of Nickel nanoparticles. Then precipitate is taken in silica crucible and heating the precipitate in the water bath at the temperature of ~96° C to get the dried black Ni nano powders. The XRD pattern of the sample was collected with a Rigaku D/Max-2000 diffractometer equipped with a Cu Ka radiation source (l ~ 0.15418 nm).

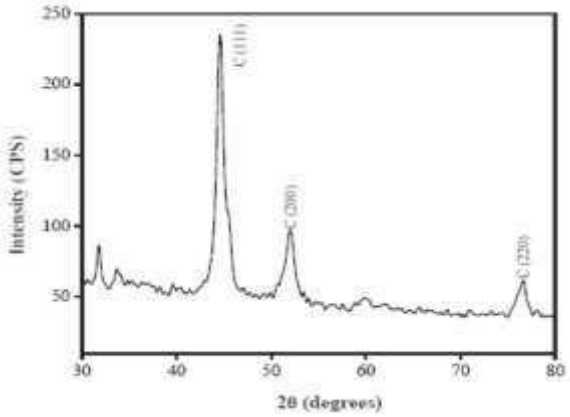
II. RESULTS AND DISCUSSION

3.1 Structural Studies

Fig1 show a typical XRD pattern of the Nickel sample. From the Figure three characteristic peaks for Nickel [2θ = 44.6˚, 51.9˚, 76.4˚] corresponding to Miller indices (111), (200), (220) were observed. This revealed that the resultant particles were pure face-centered cubic (FCC) Nickel. No obvious peaks of Nickel oxides and hydroxides were detected, possibly attributed to the observed phenomenon that Nitrogen gas was produced and bubbled up continuously during the reaction. It could be suggested that the above said gas produced might auto create an inert atmosphere; hence the input of extra Nitrogen gas was not necessary for the synthesis of Nickel nanoparticles [4, 5].

In Fig 1, the intensity ratio of the peaks [111]: [200]: [220] were calculated to be (100, 26, 160) which exactly coinciding with the JCPDS [no: 040850] values. This confirms the presence of Nickel particles in the sample and it is in good agreement with the earlier reports [7, 10-12]. However from the Fig 1 intensity ratios are not coinciding with the JCPDS values. This observed variation may be due to the change in the synthesis parameter (i.e., due to the heat at ~96˚C). The lattice constant and average particle sizes of the sample calculated from the full-width at medium height of the peaks in their XRD patterns using Scherrer formula (1) were summarized in the Table 1.

Figure-1: XRD Spectra of Nickel Nanoparticles



Sources: Authors Compilation.

[93] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mhlhemu@gmail.com
[94] Associate Professor, Government College of Technology, Tamil Nadu, India, nsuri22@gmail.com
[95] Assistant Professor, Tamilnadu College of Engineering, Tamil Nadu, India, prabusr80@gmail.com

Table-1: Structural Parameters of Nickel Nanoparticles

Sample	2θ Degree	FWHM (β) Degree	Hkl planes	Experimental	Standard cubic (JCPDS-040850)	Particle Size (D) (nm)	Lattice Constant a (Å)
				d (Å)	d (Å)		
	44.6	0.9444	111	2.03003	2.0340	9.503	3.5161
N_1	51.9	1.2800	200	1.76139	1.7620	7.213	2.4909
	76.5	1.2300	220	1.24417	1.2460	8.593	3.5190

Sources: Authors Compilation.

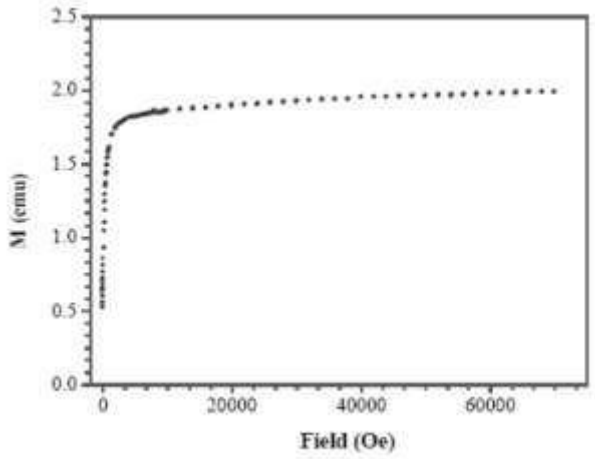
3.2 Magnetic Studies

The magnetic measurements were carried on the Nickel nanoparticles at a temperature 5K and 300k.Fig 2 shows the field dependence of the Magnetization M at 300K. Within the experimental error, hysteresis was not observed at 300K, which clearly explains the absence of Coercive force. Usually, the magnetic particles having zero coercive force (H_c) exhibits super paramagnetic state rather than ferromagnetic state [13].

This is due to the quantum size confinement effect, in Nickel nanoparticles. A gradual increase of magnetization (M) is observed below and above 2K Oe respectively. The saturation Magnetization (M_s) has been calculated from the Fig.2 as M_s=0.9236 emu/g. This is lower than the constant saturation value 55 emu/g of bulk Nickel [14].

Since the magnetization of bulk Nickel shows saturation exceeding 0.2K Oe [15], the gradual increase around 1K Oe is described to the super paramagnetic behavior. The decrease in M_s might be due to the decrease in particle size and accompanied increase in the surface area.

Figure-2: Field Dependence of Magnetization M at 300K



Sources: Authors Compilation.

The Figure3 displays the field dependence of the magnetization at 5K. Existence of hysteresis is clearly observed below 2.5K Oe. A striking feature of the hysteresis curve is that the magnetization does not saturate even at 70K Oe. It is due to the existence of paramagnetic components. This is because the particles do not have adequate thermal energy to attain complete thermal equilibrium with the applied field during the measurement time and hence exhibited a hysteretic feature.

In Fig 3 the saturation magnetization (M_s) and the remanent magnetization (M_r) for the Nickel nanoparticles at 5K are 2.55 emu/g and 1.11 emu/g, respectively. It was observed as in Figs 2 and 3 that the magnetization strongly increased on decreasing temperature. This confirms the super paramagnetic behavior of the Nickel nanoparticles.

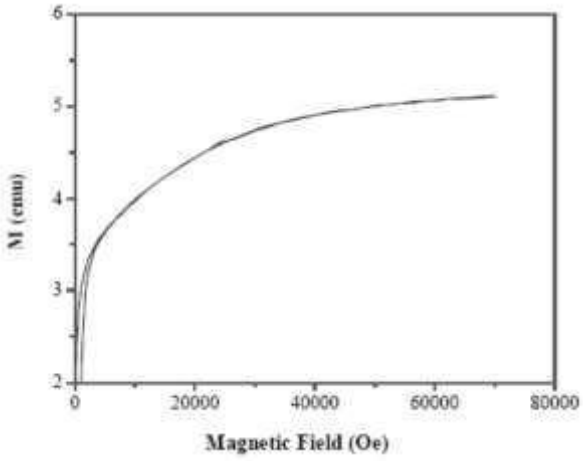
The magnetic molecules on the surface lack coordination and the spins are likewise disordered. This phenomenon is more significant for nanoparticles due to their large surface-to-volume ratio and may be another factor that leads to the decrease in M_s [16].

The lower value of saturation magnetization (Ms) compared to bulk nickel can be explained in terms of the core-shell morphology of the nanoparticles. If this surface layer were absent, the magnetization of the particles would saturate with increase in applied field, up to a particular magnetic field the core magnetic moments align with the applied magnetic field.

At some stage, resource of "core mode" of the magnetization response is exhausted and the "core magnetization" of the system is saturated in a usual Langevin – like way. Beyond this stage any increase in the magnetic field on the particles has an effect only on the surface layer of the particles and thus the increase in the magnetization of the particles slow down.

The specific state of the surface plays a virtual absence of saturation magnetization that keeps the hysteresis loops unclosed even in very strong fields. As discussed above, a protective layer from ethylene glycol might be formed on the particle surface. The protective layer also could cause the decrease in saturation magnetization (M_s). Accordingly, it is reasonable that the magnetization of nanoparticles is usually smaller than that of corresponding bulk materials [5].

Figure-3: Field dependence of Magnetization M at 5K.



Sources: Authors Compilation.

III. CONCLUSIONS

The Nickel nano particles have been synthesized by chemical reduction method. As per the XRD analysis the structure of the Nickel nanoparticles is found to be FCC crystalline in nature, with the average particle size calculated as 12 nm. The values of saturation magnetization of Nickel nano particles were found as 0.9236 emu/g and 2.55 emu/g and these of the remanent

magnetization as 0.298 and 1.11 emu/g at 300 K and 5K, respectively. These quantities exhibit an increasing trend with decrease in temperature as expected. The size reduction in Nickel nanoparticles leads to the change in ferromagnetic to super paramagnetic state.

III. REFERENCES

[1] J., H. Fendler, (1987). Chem. Rev., 87, pp. 877.

[2] G., Schmid, (1992). Chem. Rev., 92, pp.1709.

[3] L., N. Lewis, (1993). Chem. Rev., 93, pp. 2693.

[4] D., H. Chen, and S., H. Wu, (2000). "Sythesis of Nickel Nanoparticles in Water-in-Oil Microemulsions", Chem. Mater. 12, pp. 1354.

[5] Szu-Han, Wu, and Dong-Hwang, Chen, (2003). "Synthesis and Characterization of Nickel Nanoparticles by Hydrazine Reduction in Ethylene Glycol", Journal of Colloid and Interface Science, 259, pp. 282.

[6] Yang, long Hou, and Song, Gao, (2003). "Monodisperse Nickel Nanoparticles Prepared from a Monosurfactant System and their Magnetic Properties", Journal Mater. Chem., 13, pp. 1510.

[7] Xiao-Min, Ni; Xiao-Bo, Su; Zhi-Ping, Yang, and Hua-Gui Zheng, (2003). "The Preparation of Nickel Nanorods in Water-in-oil Microemulsion", Journal of Crystal Growth, 252, pp. 612.

[8] D., Kumar; S., J. Pennycook; A., Lupini; G., Duscher; A., Tiwari, and J., Narayan, "Synthesis and Atomic-Level Characterization of Ni Nanoparticles in Al_2O_3 Matrix", Appl. Phys. Lett. 81, pp. 4204.

[9] Yingwen, Duan, and Jiangong, Li, (2004). "Structure Study of Nickel Nanoparticles", Materials Chemistry and Physics, 87, pp. 452.

[10] Y., Hou, H.Kondoh, T.Ohta, S.Gao, "Size –Controlled Synthesis of Nickel Nanoparticles'', Applied surface Science 241 (2005) 218.

[11] D., E. Zhang; X-M., Ni; H., G. Zheng; Y., Li; X., J. Zhang, and Z., P. Yang, (2005, 2011). "Synthesis of Needle–Like Nickel Nanoparticles in Water-in-oil Microemulsion", Materials Letters, 59.

[12] Meining, Wu; Yong, Chun Zhu; Huagui, Zheng, and Yitai, Qian, (2002). "A Metal-Reduction Route to Synthesis of Nickel Nanotubules", Inorganic Chemistry Communications, 5, pp. 971.

[13] U., Erb; A., M. El. Sherik; G., Palumbo, and K., T. Synthesis, (1993). "Structure and Properties of Electroplated Nanocrystalline Materials", Nanostruct. Mater., 2, pp. 383.

[14] H., Danan; A., Herr, and A., J. P. Meyer, (1968). "New Determinations of Saturation Magnetization of Nickel and Iron", Journal Appl. Phys., 39, pp. 669.

[15] C., Kittel, (1996). "Introduction to Solid State Physics", 7th ed., Wiley and Sons, New York, 1996, Chapter 15, p.448.

[16] Dong-Hwang, Chen, and Chih-Hsuan, Hsieh, (2002). "Synthesis of Nickel Nanoparticles in Aqueous Cationic Surfactant Solutions", Journal Mater. Chem., 12, pp. 2412.

*****

OPTIMIZATION OF OPERATIONAL PARAMETERS FOR BALL MILL GRINDING PROCESS OF COMPOSITE CEMENT- A REVIEW

S. Charles[96] S. Irudayaraj[97]

ABSTRACT

In many cement industries the optimization of the mill grinding process for the production of composite cement i.e. Portland puzzolona cement is a very costly, time consuming task, due to many non-linear events. In India and other part of the world, the scientist are involved to increase the addition of fly ash in composite cement manufacturing, in order to make the environment more cleaner. Quality product of composite cement with higher addition of fly ash depends on the quality (hardness) of clinker, fly ash, and the particle size. Also the mill output and the cement strength is greatly influenced by addition of fly ash with clinker.

The usual problem in obtaining the desired mill output and cement strength involves selecting a set of input variables (control variables) such as specific weight of clinker, feed rate of clinker, feed rate of fly ash which will result in the desired outputs (responses) such as mill output and cement strength. Design of Experiments is one of optimization techniques which could satisfactorily be utilized for solving such problems.

The main objective of this paper is to review the literature available in this area of research.

KEYWORDS

Cement Mill, Fly Ash, Design, Experiments, Cement Clinker, Composite Cement etc.

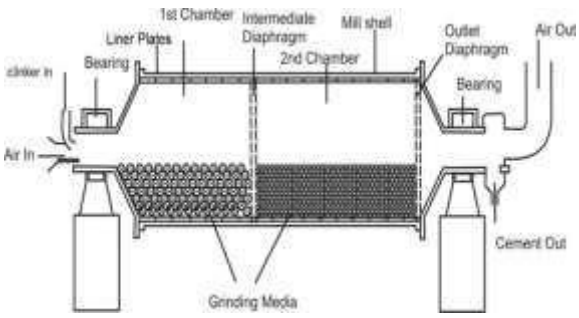
I. INTRODUCTION

1.1 Construction and Working of Ball Mill

A Ball mill is the equipment used to grind the hard, nodular clinker from the cement kiln into the fine grey powder that is cement. Most types of cement are produced using the Portland clinker and the ball mills are widely used for the grinding process of Portland clinker. A ball mill is a horizontal cylinder partly filled with steel balls (or occasionally other shapes) that rotates on its axis, imparting a tumbling and cascading action to the balls which is present inside the mill. Material fed through the mill is crushed by impact and ground by attrition between the balls which is also called as grinding media. The smaller grades of grinding media are occasionally cylindrical ("cylpebs") rather than spherical. There exists a speed of rotation (the "critical speed") at which the contents of the mill would simply ride over the roof of the mill due to centrifugal action.

The mill is usually divided into at least two chambers as shown in Fig. 1, allowing the use of different sizes of grinding media usually made up of high-chromium steel. Large balls of diameter in the range of 60-90 mm are used at the inlet, to crush clinker nodules and the media in the second chamber are typically in the range 15-40 mm, although media down to 5 mm are sometimes encountered. As a general rule, the size of grinding media has to match the size of material being ground because large media cannot produce the ultra-fine particles required and small media cannot break large clinker particles. A current of air is passed through the mill which helps to keep the mill cool, and sweeps out evaporated moisture which would otherwise cause hydration and disrupts material flow. The dusty exhaust air is cleaned usually with bag filters, which also serves as a cement dust collection system.

Figure-1: Cross Sectional View of a Cement Mill



Sources: Authors Compilation.

1.2. Grinding Systems and Mill Output

The output achieved by a ball mill system varies with the mill power, the fineness of the product and the hardness of the clinker. The hardness of clinker is an important factor that impacts the energy cost of the grinding process and it depends both on the clinker's mineral composition and its thermal history. The easiest-ground clinker mineral is alite, and so high-alite clinkers reduce grinding costs, although they are more expensive to make in the kiln.

The toughest mineral is belite, because it is harder, and is somewhat plastic, so that crystals tend to flatten rather than shatter when impacted in the mill. The mode of burning of the clinker is also important since clinker rapidly burned at the minimum temperature for combination, then rapidly cooled, contains small, defective crystals that grind easily and on the other hand, long burning at excess temperature, and slow cooling, lead to large, well-formed crystals that are hard to grind resulting in double milling costs.

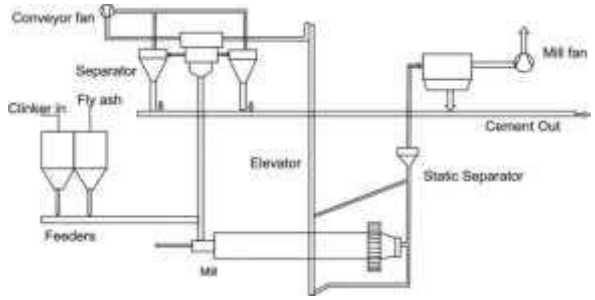
Grinding systems are either open circuit or closed circuit system [1]. In an open circuit system, the feed rate of incoming clinker and fly ash are adjusted to achieve the desired fineness of the product at the mill exit. In a closed circuit system, coarse particles are separated from the finer product and returned for further grinding. In a closed circuit mill, the total throughput is higher and hence the mill exit material is coarser.

Unlike the open circuit mill, material coming from the mill goes to an air separator from where the finer materials are the product and the rejects are returned to the mill. Fig2 presents

[96]Principal, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, chs_ani@rediffmail.com
[97] Professor, Shri Angalamman College of Engineering and Technology, Tamil Nadu, India, irudayaraj67@gmail.com

the design of the grinding system (closed circuit), which consists of a two-chamber ball mill and the separators.

Figure-2: Schematic Layout of a Cement Mill Circuit



Sources: Authors Compilation.

Grinding processes are particularly energy-consuming. In the cement industry, the energy consumption associated to grinding raw materials and clinker represents approximately 75% of the cement production costs. Moreover, only 2%–20% of the energy supplied to the grinders are effectively used in the fragmentation process, the rest being dissipated essentially into heat [2].

The use of ball mills in cement manufacturing leads to large power consumption. In addition, their capacity and operation must be optimized in order to obtain efficient performance. The performance optimization of such mills will be possible if the technological parameters of the milling process are known.
The performance of a grinding process can be influenced at two distinct levels. At first, significant improvements can be achieved by a careful design of the grinding mill (e.g. number and size of the compartments, internal lining, etc.) and the air separator (e.g. registers, velocity of the fan, etc.), and by the proper selection of grinding media (i.e. load and size distribution of the balls). Then, based on these design parameters, the development of control schemes ensuring stable process operations and efficient regulation of the cement quality is required in order to avoid product downgrading or rejection [3].

Grinding circuits are complex processes characterized by a number of design parameters and variables such as grindability of the raw materials, particle size distribution, material hold-up, and rotational speed of the mill, etc. The general lack of sensors in industrial plants makes the investigation of the several material grinding and transportation mechanisms inside the mill particularly delicate, or even impossible. The complexity of system and the difficulty to observe internal phenomena do not allow a thorough knowledge of the process. In an attempt to overcome these difficulties, numerous modeling studies of grinding processes have been carried out in the last decades.

1.3. Composite Cement (Portland Puzzolona Cement)

Fly ash is an artificial pozzolanic material, which is an inorganic residue obtained from the burning of pulverized coal and have cementitious properties, when mixed with cement is referred as composite cement. As such, the generation/ consumption ratio of fly ash is very high in India, hence every year a huge stock of fly ash is increasing. As it comes under hazardous waste material, the disposal is also a major problem. Henceforth this fly ash is used as an additive for many products and one of these products is composite cement.

Earlier, when IS1489:1969 for PPC was introduced in India, the addition of fly ash was limited from 10 – 25 % only. From 1980 onwards there is a revolutionary change in cement technology and the production of high strength giving clinker started. Based on the production of high strength clinker, Bureau of Indian Standard revised its limit and was increased and now the limit of fly ash addition in PPC is 15 – 35 % [4]. This is not only for "Waste to best" conversion but also for "ash to cash", and so it encourages such environment friendly approaches for better tomorrow. Presently OPC 43 Grade, OPC 53 Grade and PPC (Fly Ash based) cements are being manufactured and marketed.

II. LITERATURE REVIEW

Grinding technology has been continuously improving with numerous innovations with a view to improve productivity and reduce power consumption. In order to reduce the manufacturing costs for cement, it is very important to optimize the existing mill installations as far as the grinding process is concerned and also to use high quality spare parts and consumables like grinding media.

Ball mills are predominantly used machines for grinding in the cement industry. Although ball mills have been used for more than one hundred years, the design is still being improved in order to reduce the grinding costs.

The need for process optimization has been increasingly felt as production costs are shooting upwards in conjunction with increased competition in the market. In order to reduce the manufacturing costs for cement, it is very important to optimize the existing grinding installations.

Experience has shown that the potential for optimization is greatest in the cement grinding process in a cement plant. The benefits that can be achieved due to the optimization of cement grinding system are:

- Reduction in specific power consumption
- Increase in production
- Stable and sustained operation
- Increase in availability of the equipment – Less down time
- Improved and consistent product quality

Mumuni, et. al., [1] experimented the Radiotracer Residence Time Distribution (RTD) method to investigate the process of clinker grinding in Ghana Cement Plant (GHACEM) at Tema with the objective of determining hold-up and grinding efficiencies of two ball mills operating in close circuit regime. The experiment was conducted using 40Ci Au-198 radiotracer in liquid state and highly sensitive NaI detectors for radiation measurement. The experimental RTD data revealed that the Mean Residence Times (MRT) of the material in the milling and separator sections of both mills were the same. It was also observed from the estimated mill efficiencies that mill 4 operated with optimal performance while the efficiency of mill 3 was far below the expected value.

Michael Boulvin, Alain Vande Wouwer [3] has presented a nonlinear distributed parameter model of a grinding circuit used in the cement industry. The several unknown model

parameters are estimated from experimental data collected on the real plant. On this basis, a dynamic simulator is developed, which can be used to analyze the system dynamics and to explain experimental observations of real-life operation. The nonlinear dependency of the rates of breakage on the mill hold-up, and the nonlinear dependency of the separation curve on the mill flow rate have a significant influence on the process behavior, and are at the origin of a bell shaped functional relationship between the mill flow rate and material hold-up. Hence, it is required to develop a control strategy to stabilize the process operation close to the optimal point, where gain changes sign. A cascaded control structure is designed for regulating the mill flow rate, and a simple PI control loop is proposed for regulating the cement fineness. Experimental results demonstrate the good performance of the proposed control scheme. If online measurements of the recirculated flow rate are available online, then an alternative control structure, based on a subcompensation of the variations of the feed flow rate can also be considered.

For the ball mill circuit of cement mill, a high precision sampling fuzzy logic controller with self-optimizing algorithm is designed and discussed by Hui Cao et.al. [5]. The classical controllers, such as Proportional Integration Differential (PID), are of no effect. Recently many advanced control methods for the ball mill circuit have been introduced and discussed such as Nonlinear learning control method [6] and predictive control[7,8]. Oguz h.Dagci.M.Onder efe and Okyay kaynak [6] investigated a nonlinear learning control technique which is based on VSS theory, which is well known with its robustness to unmodelled dynamics and disturbances. The proposed scheme has been tested on a validated mathematical model of a cement milling circuit, which is highly nonlinear, and which has three state variables with two control inputs. The results presented demonstrate that the proposed technique is useful in control of such complex systems both in the sense of tracking accuracy and in the sense of robustness and computational simplicity.

However, these control methods mainly depend on the model of ball mill circuit [9] which is so complex since some parameters cannot be measured online. In addition, the model is not fit for every type of ball mill. The neural network control method [10] has been introduced and it includes the training algorithm that makes the work of designing the controller difficult. Fuzzy control does not need the mathematical model of the plant in addition to that of the ball mill. A controller based on fuzzy rules [11] is introduced in which the control rules are obtained from the knowledge of experts and operators. As the accumulation of errors cannot be estimated easily, the analysis and manipulation of experts and operators are usually based on the error and the error change of a controlled variable, which are usually the input variables of the general fuzzy controller.

Based on fuzzy control theory, designing the controller attains greater facility without working out the cement ball mill model and this controller becomes more practical. Based on the sampling control strategy, this controller can solve the problem of the large delay time which is the difficulty of the cement ball mill load control. With the self-optimizing algorithm, this controller can find the optimal set value itself and maintain it, so it can be better modified with the change of controlled plant. Moreover, the fuzzy interpolation algorithm allows the controller to be more sensitive, and simulation results also verify that the high precision fuzzy controller has less steady state error, compared with the general fuzzy controller.

Furthermore, the field service result also demonstrates that the controller works well in a cement mill. This controller can be applied in other situations, including ball mill circuits, such as Turbine Governor Systems of Thermal Power Plants and mineral processing of mines. In addition, the fuzzy interpolation algorithm, which is an improved method of fuzzy control, the sampling control strategy and the self-optimizing algorithm, is useful in control of such complex industry systems.

Hanifi Binici et. al. [12] explored the capability of Genetic Expression Programming (GEP) in prediction of heat of hydration of Portland and blended cements. Data for development of the models were obtained from the experimental studies.

Two GEP models were developed to generate the two mathematical equations which are contain the different number of independent variables. The results obtained from two GEP models developed herein are in satisfactory agreement with the experimental results. They have concluded that the proposed GEP models can be used as an important decision support mechanism to assist technicians in the prediction of heat of hydration of Portland and blended cements.

The introduction of modern control methodologies in traditional process industries is in steady growth. Model-based controllers usually allow to tackle multivariable control where classical PID controllers would only apply to single input single output feedback loops. In addition, recent nonlinear control techniques also provide additional improvements by allowing processes to be run over a larger operating range without controller retuning. Lalo Magni, Georges Bastin, and Vincent Wertz [7] have investigated on the application of model-based nonlinear multivariable control to a cement milling circuit. Traditionally, the application of feedback control techniques to cement milling circuits is limited to mono-variable classical proportional integral (PI) control of the tailings flow rate with either the feed flow rate or the classifier speed as control action.

A typical example has been discussed by C. Ciganek and K. Kreysa [13].Recently multivariable control techniques (based on the linear quadratic control theory) have been introduced to improve the performances of the milling circuit [14]. However, this controller, whose design is based on a linear approximation of the process, is only effective in a limited range around the nominal operating conditions. On some occasions, it has been observed on real plants that important changes in operating conditions (changes in the hardness of the clinker) have driven the mill to a region where the controller cannot stabilize the plant.

A. Seyfi erdem, S. Levent ergun, A. Hakan benzer [15] derived a new approach based on Morrell's C model is used to calculate the power draw of dry multi-compartment ball mills. Calculated power draws were in good agreement with the measured values. Volumetric mill load calculated using different equations gave similar results. Increasing the number of measurements taken along the width and length of the mill would improve the accuracy of the calculation. It was found that the power draw of dry multi-compartment ball mills used in cement grinding could successfully be predicted using this approach. However, this method needs to be validated with

more data sets including variation of other operating parameters such as critical speed and lifter design. Grinding of cement clinker has been traditionally performed by the ball mills which can either be open circuit or closed circuit with an air classifier. Selection of right mill for the specified duty is the most critical for circuit design, since it has the highest capital and operating costs. Bond method has been used for ball mill selection in both mineral and cement industry for 50 years. It is basically rely on determination of grindability of material in a specified laboratory mill. Then, using empirical equations developed by Bond, and later with minor revisions by Rowland, the mill size is determined [16,17].

Once, the size and ball load is determined, the power draw for a given mill is calculated. A multi-compartment ball mill consists of two or more grate discharge ball mills in series. The same equation is used to calculate the power that each ball mill compartment should draw. The total power is the sum of the power calculated for each of the separate compartments. Although, Bond's method has been widely used, the required link to classification is missing. Therefore, it is not possible to calculate circuit performance when any of the design and operational parameters are changed [17].

Model based methods has becoming popular to overcome these deficiencies[18-21].Morrell and Man [21] proposed a model based approach using the results of the Bond ball mill grindability test for overflow wet ball mills. For all model based methods, a reliable method to calculate mill power draw for a given mill is required for the calculation of power draw. Morrell proposed a mathematical model for autogenous, semi-autogenous ball mills which is based on the motion of grinding charge inside the mill and also he verified his approach with various plant data [17].

Andon Venelinov Topalov, Okyay Kaynak [22] proposed a novel neural network adaptive control scheme for cement milling circuits. Its major advantage is that there is no need to develop an accurate model of the milling circuit in advance. Instead, it is learned online by a neural network. This allows dealing more successfully with the existing model uncertainties, parameter changes and their interrelations. The possibilities for implementation of a sliding mode learning algorithm, proposed as an on-line mechanism for adaptation of neuro control systems have been investigated. It is confirmed that the algorithm can be used to train the networks as they interact with the external environment. The trained structures are robust and learn fast, both of which are features inherited from SMC.

Simulations results show that the nominal responses to the set-point changes as well as to the hardness change are correct and these changes do not cause instabilities of the closed loop system (plugging). In can be observed also that the transition times, undershoots and overshoots are smaller .The real difficulty related to the practical implementation of the proposed control approach is the need for a measurement of the load of the mill for which the cement mill could be installed on a weighing system and its total mass can be measured by which the cement load can be derived. Some mills are equipped with an ''electronic ear'' wherein the measured noise is inversely proportional to the mill load.

Javed I. Bhatty, John Gajda, and F. M. Miller [23] performed two commercial-scale demonstrations of the use of high-carbon fly ash in cement manufacture and the results from the initial demonstration have proven three critical benefits. First, new market for high carbon fly ash has been developed. High carbon fly ash, when used in this regard requires no preprocessing. Second, the fly ash replaces costly raw materials such as shale and clay normally used in as raw materials in cement manufacture.

Finally, the high carbon content conserves energy by serving as a partial fuel substitute in the energy intensive cement manufacturing operation. Nearly 65 million tons fly ash is annually produced in the U.S., of which only one third is used in commercial products. The majority of the fly ash that is generated is landfilled. With the impending regulatory pressure to reduce emissions at coal-fired power plants, production of more fly ash with significantly higher carbon contents is imminent. Since high carbon fly ashes are undesirable for use in concrete, the urgency for alternate reuse applications is growing.

In each of the highly successful demonstrations, a large quantity of high-carbon fly ash was procured from Ameren's Power Station in Coffeen, Illinois, and used in a nearby cement plant. The cement plant utilizes preheater kiln system. Prior to its use, the ash was analyzed for chemical composition and evaluated for its compatibility with raw materials from the cement plant. The analysis of the fly ash confirmed its compatibility with the normally used raw materials. At the cement plant, the fly ash was inter-ground with the other raw materials and introduced into the manufacturing process in a normal manner. The raw materials were then normally processed in to clinker (an intermediate product).

The resulting clinker was ground in the finish mills with gypsum to produce portland cement. Both the clinker and cement were characterized and compared to those normally produced at the plant. During the demonstration several key parameters on the aspect of material processing, operation, and product performance were also observed. Although the intent was to maximize the use of fly ash in the raw mix, fluctuations in the chemistry of the other raw materials limited the average addition of fly ash to 6% of the raw mix. The fly ash, however, replaced a majority of the shale in the raw mix.

During the preheating stage, the temperatures of the lower preheater stages rose significantly primarily because of the thermal contribution of carbon in the fly ash. This resulted in nearly 30% additional calcination of the raw feed. Typically, an improved calcination facilitates burning, conserves fuel, and improves clinker production. Also, the material flow within the preheater cyclones was smooth and plug-free, because of the reduced exit pressures. An increase in kiln feed rate was also noted which can be attributed to high carbon that had caused an increased calcination of the raw feed. Increased kiln feed rate resulted in 10% improvement in clinker production.

Furthermore, fuel supply during the demonstration was also reduced to adjust for the energy provided by carbon in the fly ash and the resulting energy reduction was nearly 4%.Characterization of the demonstration clinker confirmed the presence of major clinker phases. Of interest is the reduced alkali content of cement without any equipment modifications.

Low alkali cements are always preferred in concrete for improved durability. The compressive strength of cement was comparable to cement produced before to and after the demonstration. In fact the 28-days strength for cement made

during the demonstration improved, which can potentially produce higher strength concrete. This technology offers a large-scale consumption of high-carbon fly ash that can translate to improved economics to the utilities as well as to the cement manufacturers.

Clinker formation in coal fired rotary cement kilns under realistic operation conditions has been modeled by E. Mastorakos et. al. [24] with special emphasis given to the heat exchange between the coal flame and the kiln, the heat exchange between the kiln and the counter flowing solids, and the chemical reactions that transform the solids into the final product (the clinker). An axisymmetric CFD code (the commercial package FLOW-3D) that includes a Monte-Carlo method for radiation has been used for the gaseous phase. The temperature of the kiln wall has been calculated with a finite volume heat conduction code, and the species and energy conservation equations for the clinker have also been formulated and solved.

An iterative procedure between the predictions for the gaseous temperature, the radiative heat flux to the wall, and the kiln and clinker temperatures was used to calculate the distribution of the inner wall temperature explicitly, in contrast to previous modeling efforts that used such information from experiments. Integration over the tangential coordinate resulted in an axial distribution of an "effective"' inner wall temperature, used as the necessary boundary condition to the CFD code. This procedure makes the calculation of heat flow to the clinker possible, and also allows the approximate treatment of a 3-D problem with an axisymmetric CFD code.

The results show that radiation accounts for most of the heat transfer between the gas and the kiln walls, while the heat loss through the refractories to the environment accounts for about 10% of the heat input. The chemical reactions and heating of the charge absorb about 40% of the energy of combustion; it is hence imperative that these are accounted for. The predictions are consistent with trends based on experience and available measurements from a full-scale cement kiln.

Cement Kiln Dust (CKD) produced in a local cement production plant in Saudi Arabia along with fly ash resulting from combustion of heavy fuel oil in a local power generation plant were utilized by M. A. Daous [25] as waste material blends with Portland cement, produced from the local plant, at various proportions. The following conclusions can be made concerning the performance of such blends:

(1) Mortars of satisfactory mechanical strength can still be produced using blends containing 90% Portland cement but not more than 4% fly ash as blended waste material.
(2) Blends containing as low as 70% Portland cement can still exhibit adequate strength if only CKD is used as the blending waste material.
(3) These blends witness various degrees of increase in their hydration requirements for normal consistency, and more noticeably in their initial setting times.

Various methods for utilizing CKD in industrial applications, including existing or proposed methods for alkali removal, are reported in the literature.

Bhatty [26] provides a general review of these methods. Because of the generally high lime content of CKD and subsequent ability to harden upon exposure to moisture, CKD has been used as a binder in soil stabilization suitable for a sub base in streets and high way construction. It is commonly used as a mixture with different solid- waste materials such as waste glass, fly ash, waste water sludge with the addition of cement or other admixtures if necessary [27,28]. Alkalis may be removed from CKD by aqueous leaching followed by returning the washed CKD to the kiln [29, 30], or by fluidized bed technology to volatilize alkalis.

A number of patents have been issued dealing with the extraction of soluble potassium salts that are later recovered as a saleable product and the water recovered for reuse. The potassium-free dust is usually recycled as a kiln feed. A great deal of work has been done on the use of CKD in blended cements.

Bhatty has published a series of reports [31-34] on the addition CKD Portland cement along with fly ash and blast furnace slag, with variable results. Cements blended with CKD alone reportedly had reduced strength, setting time, and workability. The strength loss was attributed to the presence of high alkalis in the dust. The addition of fly ash with CKD lowered the alkalis contents and resulted in improved strength. Additions of slag generally reduced workability but improved blends' strengths because of activation from the high lime content of CKD. Bhatty's results also demonstrated that particular ratios of alkalis, chlorides, and sulfates are important for better performance of cement blends.

A high sulfate blend, for instance, gave a high strength. However, high alkali contents results into increased susceptibility to alkali – aggregate reaction, which may be blocked by the slag or fly ash constituents. The researchers found that adding CKD to blends of ordinary Portland cement, blast furnace slag cement, or sulfate resistant cement beyond specific upper limits to each of these blends adversely affected the physical and mechanical properties of these blended cement pastes.

The utilization of different types and classes of fly ash as a partial replacement of cement or aggregates in concrete mixes or as an additive in different blends of cement mortars has been investigated by many researchers in the last few years. It has been reported that such utilization produces various technical, economical, and environmental benefits [35-37]. These studies also show that the resulting mechanical properties of blended mortars depend to a large extent on type of fly ash used, its source, fuel type, combustion temperature, particle size, and other factors.

III. DESIGN OF EXPERIMENT (DOE)

Design of Experiment is an optimization technique which involves designing a set of experiments, in which all relevant factors are varied systematically. When the results of these experiments are analyzed, they help to identify optimal conditions, the factors that most influence the results, and those that do not, as well as details such as the existence of interactions and synergies between factors. DOE methods need well-structured data matrices and when applied to a well-structured matrix, analysis of variance delivers accurate results, even when the matrix that is analyzed is quite small. Experimental design is a strategy to gather empirical knowledge, i.e. knowledge based on the analysis of experimental data and not on theoretical models. It can be

applied whenever you intend to investigate a phenomenon in order to gain understanding or improve performance.

Building a design means, carefully choosing a small number of experiments that are to be performed under controlled conditions. There are four interrelated steps in building a design:

1. Define an objective to the investigation, e.g. better understand or sort out important variables or find optimum.
2. Define the parameters that will be controlled during the experiment (input variables), and their levels or ranges of variation.
3. Define the parameters that will be measured to describe the outcome of the experimental runs (responses), and examine their precision.
4. Among the available standard designs, choose the one that is compatible with the objective, number of design variables and precision of measurements, and has a reasonable cost.

Standard designs are well-known classes of experimental designs. They can be generated automatically as soon as you have decided on the objective, the number and nature of design variables, the nature of the responses and the number of experimental runs you can afford. Generating such a design will provide you with a list of all experiments you must perform, to gather enough information for your purposes.

Design of Experiments (DOE) is widely used in research and development, where a large proportion of the resources go towards solving optimization problems. The key to minimizing optimization costs is to conduct as few experiments as possible. DOE requires only a small set of experiments and thus helps to reduce costs.

IV. CONCLUSIONS

In the recent past, for the ball mill circuit of cement mill, a high precision sampling fuzzy logic controller with self-optimizing algorithm has been designed and discussed. The capability of Genetic Expression Programming (GEP) in prediction of heat of hydration of Portland and blended cements is being understood. There are investigations on the application of model-based nonlinear multivariable control to a cement milling circuit. The advantage of using a novel neural network adaptive control scheme for cement milling circuits is also present. In these literature the optimization of the operating parameters such as feed rate of clinker and fly ash, specific weight of clinker, air flow rate by utilizing the optimization techniques such as design of experiments, genetic algorithm etc. is insignificant which a potential area for research work.

REFERENCES

[1]. I., Mumuni; C., P., K. Dagadu; K., A. Danso; P., S. Adu; H., A. Affum, I., Lawson, G., K. Appiah, A., Coleman, and M., A. Addo, (2011). "Radiotracer Investigation of Clinker Grinding Mills for Cement Production at Ghacem", Research Journal of Applied Sciences, Engineering and Technology, 3(1), pp. 26-31.

[2] O., Labahn, and B., Kohlhaas, (1983). Cement Engineers' Handbook: Bauverlag, pp. 239–241.

[3] Michael, Boulvin; Alain, Vande Wouwer; Renato, Lepore; Christine, Renotte, and Marcel, Remy, (2003). "Modeling and Control of Cement Grinding Processes", IEEE Transactions on Control Systems Technology, Volume: 11, Number: 5.

[4] S., B. Roongta; T., S. Dewangan, and Usha, Singh, "Addition of Fly-Ash beyond BIS limit (as mentioned in IS 1489:1991), Portland Pozzolana Cement, A R&D work for cleaner environment".

[5]. Hui, Cao; Gangquan, Si; Yanbin, Zhang; Xikui, Ma, and Jingcheng, Wang, (2008). "Load Control of Ball Mill by a High Precision Sampling Fuzzy Logic Controller with Self-Optimizing", Asian Journal of Control, Volume: 10, Number: 6, pp. 621-631.

[6] Oguz, H. D.; M., Onder Efe, and O., Kaynak, (2001). "A Nonlinear Learning Control Approach for a Cement Milling Process", Proceedings 2002 IEEE International Confernce Control Appl., Mexico, pp. 498–503.

[7] Lalo, Magni; Georges, Bastin, and Vincent, Wertz, (1999). "Multi Variable Nonlinear Predictive Control of Cement Mills", IEEE Transactions on Control Systems Technology, Volume: 7, Number: 4.

[8] Martin, G., and S., Mcgarel, (2001). "Nonlinear Mill Control", ISA Trans., Volume: 40, Issue: 4, pp. 369–379.

[9] Ja, F.; G., Bastin; V., Wertz, and L., Magni, (1997). "Preventing Cement Mills from Plugging by Robust State Feedback Stabilisation", Proceedings Confernce Control 97, Sydney, Australia, pp. 248– 251.

[10] Yildiran, U., and M., Okyay Kaynak, (2002). "Neural Network Based Control of a Cement Mill by Means of a VSS Based Training Algorithm", Proceeding 2002 IEEE Int. Symp. Ind. Electron. L'aquila, Italy, Volume: 1, pp. 326–331.

[11] Wei, W.; Z., Yanbin; P., Haifeng, and J., Lixin, (2002). "Material Blending and Pulverization Control System in Cement Industries", Proceeding 2002 Int. Fieldbus / Control Manage. Integration Conf. Exhibition, Shanghai, China, 2002, Volume: 1, pp. 230–231.

[12] Hanifi, Binici; Cafer, Kayadelen; Gsmail, H.ÇaFAtay; Mustafa, Tokyay, and Hasan, Kaplan, (2009). "Genetic Expression Programming for Prediction of Heat of Hydration of the Blended Cements", Scientific Research and Essay, Volume: 4 (3), pp. 141-151.

[13] C., Ciganek, and K., Kreysa, (1991). "Two-Parameter Control System for a Cement Grinding Process", Translation of Zement-Kalk-Gips, pp. 202–206.

[14] V., Van Breusegen; L., Chen; V., Werbrouck; G., Bastin, and V., Wertz, (1994). "Multivariable Linear Quadratic Control of a Cement Mill: An Industrial Application", Contr. Eng. Practice, Volume: 2, pp. 605–611.

[15] A., Seyfi Erdem; Ş., Levent ergün, and A., Hakan Benzer (2004). "Calculation of the Power Draw of Dry Multi–Compartment Ball Mills", Physicochemical Problems of Mineral Processing, 38, pp. 221-230.

[16] Rowland, C. A., (1982). "Selection of Rod Mills, Ball Mills, Pebble Mills and Regrind Mills, Design and Installation of Comminution Circuits, American Institute of Mining', Metallurgical and Petroleum Engineers Inc, New Jersey, Chapter 23, pp. 393-438.

[17] Napier-Munn, T. J.; Morrell, S.; Morrison, R. D., and Kojovic, T., (1996). "Mineral Comminution Circuits their Operation and optimization", 1996, 1-413.

[18]. Herbst, J. A., and Fuerstenau, D. W., (1980). "Scale-up Procedure for Continuous Grinding Mill Design using Population Balance Models", International Journal of Mineral Processing, 7, pp. 1- 31.

[19] Austin, L. G.; Klimpel. R. R.; Luckie, P. T., and Rogers, R. S. C., (1982), "Simulation of Grinding Circuits for Design, Design and Installation of Comminution Circuits", American Institute of Mining, Metallurgical and Petroleum Engineers Inc, New Jersey, Chapter 19, pp. 301-324.

[20] Austin, L. G.; Klimpel, R. R., and Luckie, P. T., (1984). "Process Engineering of Size Reduction: Ball Milling", American Institute Of Mining, Metallurgical and Petroleum Engineers Inc, New Jersey, pp. 1-556.

[21] Morrell, S., and Man, Y. T., (1997), "Using Modelling and Simulation for the Design of Full Scale Ball Mill Circuits", Minerals Engineering, Volume 10, Number; 12, pp. 1311-1327.

[22] Andon, Venelinov Topalov, and Okyay, Kaynak, (2004). "Neural Network Modeling and Control of Cement Mills using a Variable Structure Systems Theory Based on-line Learning Mechanism", Journal of Process Control, 14, pp. 581–589.

[23] Javed, I. Bhatty; John, Gajda, and F., M. Miller, (2003). "Commercial Use of High-Carbon Fly Ash in Cement Manufacturing", International Ash Utilization Symposium.

[24] E., Mastorakos; A., Massias1; C., D. Tsakiroglou; D., A. Goussis; V., N. Burganos, and A., C. Payatakes, (1999). "CFD Predictions for Cement Kilns Including Flame Modeling, Heat Transfer and Clinker Chemistry", Applied Mathematical Modeling, 23, pp. 55-76.

[25] M., A. Daous, (2004). "Utilization of Cement Kiln Dust and Fly Ash in Cement Blends in Saudi Arabia", JKAU: Eng. Sci., Volume: 15, Number: 1, pp. 47-61.

[26] Bhatty, J. I., (1995). "Alternative Uses of Cement Kiln Dust, RP 327", Portland Cement Association, Skokie, Illinois, U.S.A. 1995.

[27] Miller, C. T.; Bensch, D. G., and Colony, D. C., (1980). "Use of Cement Kiln Dust and Fly Ash in Pozzolanic Concrete Base Courses, Emulsion Mix Design, Stabilization, and Compaction", Transportation Research Record 754, TRB, Washington D. C., U.S.A., pp. 36-41.

[28] Bhatty, M.S.Y., (1983). "Use of Kiln Dust in Blended Cements', PCA R&D Serial No.1717, Portland Cement Association, Skokie, Illinois, U.S.A.

[29] Klemm, W. A., (1993). "Cement Kiln Dust: A look at Its Uses and Characteristics", Proceedings of 29th International Cement Seminar, Rock Product, San Francisco, California, U.S.A.

[30] Young, G. L., and Morrison, G. L., (1991). "Passamaquoddy Technology Recovery Scrubber at Dragon Products Company, Inc. Cement Plant, Rock Products", 27th International Cement Seminar, Atlanta, Georgia, U.S.A.

[31] Bhatty, M.S.Y., (1984). "Use of Cement Kiln Dust in Blended Cements", World Cement Technology, London, U.K., Volume: 15, Number: 4, pp. 126-134.

[32] Bhatty, M.S.Y., (1984). "Kiln Dust Cement Blends, Evaluated, Rock Products", Chicago, Illinois, U.S.A., Volume: 88, Number: 10, pp. 47-65.

[33] Bhatty, M.S.Y., (1984). "Use of Cement Kiln Dust in Blended Cements: Alkali - Aggregate Reaction Expansion", World Cement Technology, London, U. K., Volume: 16, Number: 10, pp. 386-392.

[34] Bhatty, M.S.Y., (1986). "Properties of Blended Cement Made with Portland Cement, Cement Kiln Dust, Fly Ash, and Slag", 8th International Congress on Chemistry of Cement, Rio de Janeiro, Brazil, Theme 3, Volume: IV, pp. 118-127.

[35] Papadakis, V. G., (1999). "Effect of Fly Ash on Portland Cement Systems, Part I Low-Calcium Fly Ash", Cement and Concrete Research, Volume: 29, pp. 1727-1736.

[36] Erdogdu, K., and Turker, P., (1998). "Effects of Fly Ash Particle Size on Strength of Portland Cement Fly Ash Mortars", Cement and Concrete Research, Volume: 28, pp. 1217-1222.

[37] Bouzoubaa, N.; Zhang, M. H., and Malhotra, V. M., (2001). "Mechanical Properties and Durability of Concrete Made with High-Volume Fly Ash Blended Cements using a Coarse Fly Ash", Cement and Concrete Research, Volume: 30, pp. 1393-1402.

*****

AUTOMATIC DETECTION OF GLAUCOMA DISEASE IN EYE

S. Sangeetha[98] M. Thanga kavitha[99]

ABSTRACT

Glaucoma arises due to the inadequate fluid flow from the drainage canals of the eye, leading to the crystallization of the fluid in the cornea and iris regions. Especially in closed angled Glaucoma, fluid pressure in the eye increases because of inadequate fluid flow between the iris and the cornea. One important technique to assess patients at risk of Glaucoma is to analyze ultrasound images of the eye to detect the structural changes. In this paper, an algorithm is proposed to automatically compute this accretion from the ultrasound images of the eye.

The algorithm aims to determine the exact location of the apex point of the anterior chamber region for efficient angle calculation. It is highly imperative to detect Glaucoma in its early stages for diagnosis and hence the algorithm also addresses the importance of precise results with effective immunity towards speckle noise.

This work shows a technique to improve the efficiency of clinical interpretation of Glaucoma in ultrasound images of the eye.

KEYWORDS

Glaucoma, Drainage Canals, Cornea, IRIS Region, and Ultrasound Images etc.

I. INTRODUCTION

Glaucoma is a group of diseases that can steal sight without warning or symptoms. Some of the alarming facts about Glaucoma are (1) Glaucoma is a leading cause of blindness, (2) There is no cure for Glaucoma yet, (3) Everyone is at risk and (4) There may be no symptoms. Nearly half of those with Glaucoma do not know they have the disease. This has been shown repeatedly in studies conducted in developed countries.

Glaucoma is a potentially blinding disease that affects 66 million persons worldwide. It is the second leading cause of blindness worldwide. The disease is characterized by typical changes in the optic nerve (the nerve that connects the eye to the brain) with associated visual field defects (the area seen by the eye). Since the outer portion of the visual field is the first to be affected and most types of Glaucoma are asymptomatic the disease is often diagnosed once significant vision/field has been lost.

Therefore, early diagnosis is essential so that treatment to halt/slow progression can be instituted. Glaucoma study from Chennai city and rural Tamilnadu reveals that every 1 on 3 patients above 40 years having vision related problems were diagnosed with Glaucoma. When the Glaucoma is understood and managed, humans can continue to live their life fully.

[98]Student, Akshaya College of Engineering, Tamil Nadu, India, sangeethashanmugam09@gmail.com
[99]Student, Akshaya College of Engineering, Tamil Nadu, India, irudayaraj67@gmail.com

Glaucoma is a group of diseases of the optic nerve involving loss of retinal ganglion cells in a characteristic pattern of optic neuropathy. Eye has pressure just like blood, and when this intraocular pressure (IOP) increases to dangerous levels, it damages the optic nerve. This can result in decreased peripheral vision and, eventually, blindness.

Glaucoma is similar to ocular hypertension but with accompanying optic nerve damage and vision loss. Although raised intraocular pressure is a significant risk factor for developing Glaucoma, there is no set threshold for intraocular pressure that causes Glaucoma.

One person may develop nerve damage at a relatively low pressure, while another person may have high eye pressures for years and yet never develop damage. Untreated Glaucoma leads to permanent damage of the optic nerve and resultant visual field loss, which can progress to blindness. In this paper, the Closed Angle Glaucoma is addressed, in which the fluid at the front of the eye cannot reach the angle and leave the eye.

The angle gets blocked by part of the iris. People with this type of Glaucoma have a sudden increase in eye pressure. Symptoms include severe pain and nausea, as well as redness of the eye and blurred vision. This status requires immediate medical attention.

II. CURRENT TECHNIQUES TO DETECT GLAUCOMA

Regular Glaucoma check-ups include two routine eye tests: Tonometry and Ophthalmoscopy.

2.1 Tonometry

The Tonometry test measures the inner pressure of the eye. Usually drops are used to numb the eye. Then the doctor or technician will use a special device, called Tonometer, which measures the eye's pressure. The normal range of this pressure is in between 10mmHg and 22mmHG.

2.2 Ophthalmoscopy

Ophthalmoscopy is used to examine the inside of the eye, especially the optic nerve. In a darkened room, the doctor will magnify the eye by using an ophthalmoscope (an instrument with a small light on the end). This helps the doctor look at the shape and color of the optic nerve.If the pressure in the eye is not in the normal range (10mmHg to 22mmHg), or if the optic nerve looks unusual, then one or two special Glaucoma tests will be done. These two tests are called Perimetry and Gonioscopy.

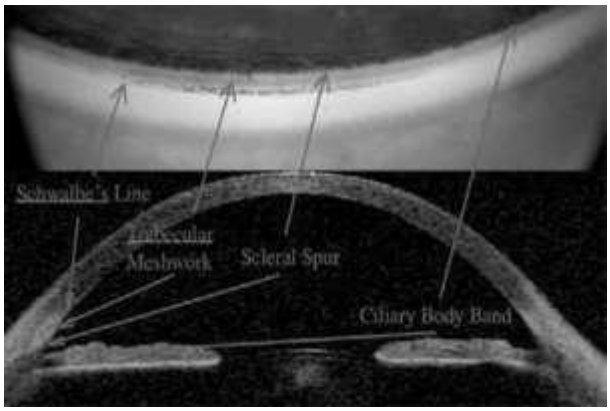
2.2.1 Perimetry

The Perimetry test is also called a visual field test. During this test, the patient will be asked to look straight ahead and then indicate when a moving light passes his/her peripheral (or side) vision. This helps draw a ―map‖ of patient's vision.

2.2.2 Gonioscopy

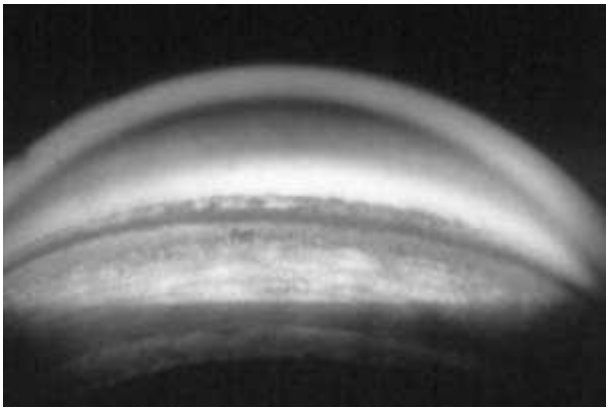
Gonioscopy is a painless eye test that checks if the angle where the iris meets the cornea is open or closed, showing if either open angle or closed angle Glaucoma is present.

Figure-2.1: Clinical Parameters in Gonioscopic Images



Sources: Authors Compilation.

Figure-2.2



Sources: Authors Compilation.

2.3 Manual Calculation of AOD from Ultrasound Images of the Eye

In this method, the doctor examines the Ultrasound image of the patient's eye and he/she estimates the angle between the iris and the cornea. If the estimated angle is less than 190, then the eye is treated as Glaucoma affected, otherwise as normal eye. But the manual estimation may be wrong.

2.4 Disadvantages of Current Techniques

Manual analysis of eye images is fairly time consuming, and the accuracy of parameter measurements varies between experts. For this reason, an algorithm is developed to automatically analyze eye ultrasound images. The proposed algorithm is expected to reduce the processing time taken by the existing techniques of manual/computer-based algorithms without compromising on the speed, accuracy, sensitivity, cost and compatibility of the product. Ultrasound images of eye are usually associated with poor resolution, poor contrast, noise and divaricated anterior chamber edges. Algorithm is proposed to effectively mitigate the above challenges.

III. ALGORITHM DESIGN

This algorithm describes a new method to detect features in ultrasound images, which shows good performance in detection of difficult features. The developed techniques make use of major image processing methods and fundamentals. In order to calculate the clinical parameters of interest, new region classification and segmentation techniques are developed as well as other signal processing techniques are used to locate the scleral spur. The ultrasound images of the eye are very noisy, with poor resolution and weak edge delineation, which required the development of a three-step method to overcome these challenges. The complete algorithm is shown in Figure 3.1

Figure-3.1

IV. RESULTS

The angles between iris and cornea for Ultrasound images of different patients are calculated and decision about the presence of Glaucoma is made follows.

Figure-4.1: Results for Ultrasound Image 1

Sources: Authors Compilation.

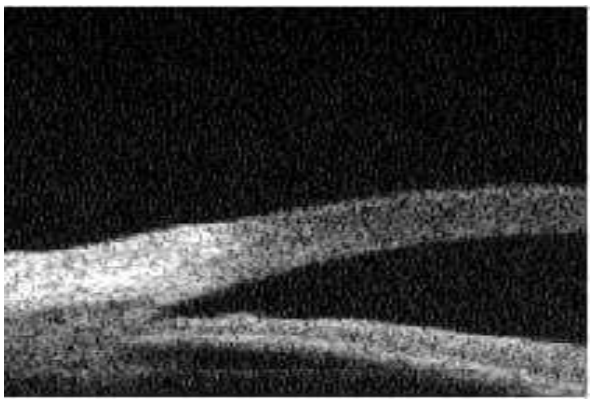
Table-1: Comparison of Results with Current Techniques for Image 1

SLNo.	Technique used	Measured with respect to	Glaucoma is present if	Result
1	Tonometry	Intraocular Pressure (IOP)	IOP > 22mmHg	IOP = 22mmHg Cannot Decide
2	Perimetry	Visual Filed of the Patient	-	Cannot Decide
3	Gonioscopy	Angle between Iris and Cornea (AOD)	AOD < 19°	AOD = 22° No Glaucoma
4	Direct view of the image	Angle between Iris and Cornea (AOD)	AOD < 19°	AOD = 25° No Glaucoma
5	Presented technique	Angle between Iris and Cornea (AOD)	AOD < 19°	AOD = 15.33° Glaucoma Detected

Sources: Authors Compilation.

As the Intraocular Pressure is same as the threshold value, the status of Glaucoma cannot be decided using Tonometry. As the Perimetry method is the visual filed of the patient, in this method also the status of Glaucoma cannot be decided. Even though AOD is greater than 190 in the case of Gonioscopy and direct view, which decided that Glaucoma is present, the developed algorithm detected Glaucoma, as AOD is less than 190.

Figure-4.2: Results for Ultrasound Image 2



Sources: Authors Compilation.

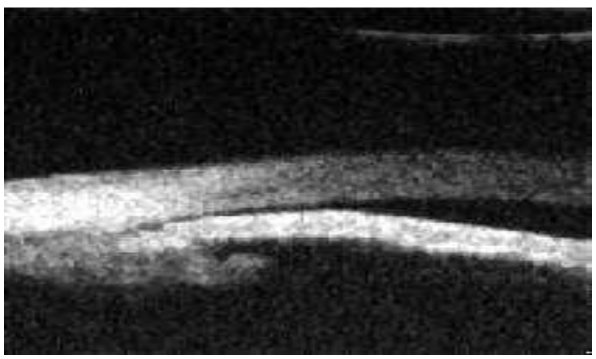
Table-2: Comparison of Results with Current Techniques for Image 2

Sl.No.	Technique used	Measured with respect to	Glaucoma is present if	Result
1	Tonometry	Intraocular Pressure (IOP)	IOP > 22mmHg	IOP = 22mmHg Cannot Decide
2	Perimetry	Visual Filed of the Patient	-	Cannot Decide
3	Gonioscopy	Angle between Iris and Cornea (AOD)	AOD < 19^0	AOD = 22^0 No Glaucoma
4	Direct view of the image	Angle between Iris and Cornea (AOD)	AOD < 19^0	AOD = 25^0 No Glaucoma
5	Presented technique	Angle between Iris and Cornea (AOD)	AOD < 19^0	AOD = 15.33^0 Glaucoma Detected

Sources: Authors Compilation.

As the Perimetry method is the visual filed of the patient, in this method also the status of Glaucoma cannot be decided. Even though Tonometry and Gonioscopy decided that Glaucoma is present, the developed algorithm detected no Glaucoma, which prevents the unnecessary surgery.

Figure-4.3: Results for Ultrasound Image 3



Sources: Authors Compilation.

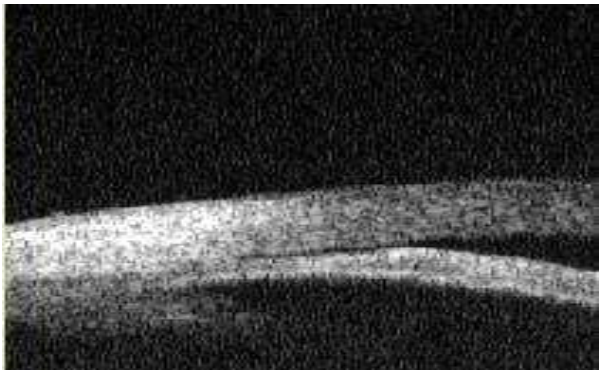
Table-3: Comparison of Results with Current Techniques for Image 3

Sl.No.	Technique used	Measured with respect to	Glaucoma is present if	Result
1	Tonometry	Intraocular Pressure (IOP)	IOP > 22mmHg	IOP = 37mmHg Glaucoma Detected
2	Perimetry	Visual Filed of the Patient	-	Glaucoma Detected
3	Gonioscopy	Angle between Iris and Cornea (AOD)	AOD < 19^0	AOD = 13^0 Glaucoma Detected
4	Direct view of the image	Angle between Iris and Cornea (AOD)	AOD < 19^0	AOD = 13^0 Glaucoma Detected
5	Presented technique	Angle between Iris and Cornea (AOD)	AOD < 19^0	Anterior Chamber almost closed (Glaucoma Detected)

Sources: Authors Compilation.

In this case, all methods including developed algorithm decided that Glaucoma is present. Perimetry method also decided that Glaucoma is present, as the visual field of the patient's eye is very poor.

Figure-4.4: Results for Ultrasound Image 4



Sources: Authors Compilation.

Table-4: Comparison of Results with Current Techniques for Image 4

Sl.No.	Technique used	Measured with respect to	Glaucoma is present if	Result
1	Tonometry	Intraocular Pressure (IOP)	IOP > 22mmHg	IOP = 35mmHg Glaucoma Detected
2	Perimetry	Visual Filed of the Patient	-	Glaucoma Detected
3	Gonioscopy	Angle between Iris and Cornea (AOD)	AOD < 19^0	AOD = 15^0 Glaucoma Detected
4	Direct view of the image	Angle between Iris and Cornea (AOD)	AOD < 19^0	AOD = 16^0 Glaucoma Detected
5	Presented technique	Angle between Iris and Cornea (AOD)	AOD < 19^0	Anterior Chamber almost closed (Glaucoma Detected)

Sources: Authors Compilation.

In this case, all methods including developed algorithm decided that Glaucoma is present.

In ultrasound imaging, speckle noise severely degrades the visual quality of the image. In order to achieve high accuracy when extracting features, speckle must be filtered without destroying any important characteristics in the image. In the developed algorithm, speckle noise was reduced using a multi-scale algorithm.

It is worthwhile to investigate a different speckle reduction technique that do not depend on the selection of the window size and that can be used on the ultrasound images of the eye before edge enhancement. One easy way to reduce speckle is to average multiple uncorrelated images of the same object obtained from different spatial positions.

However, this procedure is computationally costly and will increase the processing time of the algorithm. It seems to design an algorithm for fine enhancement, which does not require the selection of a fixed window size and to reduce speckle noise based on each pixel surrounding area.

However, for images with very poor resolution, more iteration can be applied until all pixels lying in the same local neighborhood have similar intensity values close to the initial spike value. If this technique shows improvement in speckle reduction and does not destroy edges in the original image, the enhancement process in the algorithm will require less iteration, resulting in a considerable reduction of the processing time.

V. CONCLUSIONS

This thesis has developed an algorithm to automatically identify clinical features in ultrasound images of the eye. The algorithm computes the AOD 500 used to measure the presence and severity of glaucoma. Overall, the algorithm predictions are very advantageous compared to the technologist's observation. In the processed images, features were correctly identified in 97% of the cases. 3% of images presented inaccurate approximation of the clinical parameters. The difficulties encountered in measuring clinical parameters, which are associated with the speckle noise, poor contrast, poor resolution, and weak edge delineation present in the processed ultrasound images, are accurately eliminated.

However, the designed algorithm failed for a few of images, where more noise is present. The algorithm was designed with a goal of robustness through the use of enhancement process on the original image, and by validation of the proper segmentation of the anterior chamber at each step. Overall, the benefit of this work is the ability of algorithm to reduce the processing time and improve processing consistency for each patient's ultrasound image, leading hopefully to an increase in efficiency and a reduction of cost.

VI. REFERENCES

[1] R., Youmaran; P., Dicorato; R., Munger; T., Hall, and A., Adler, (2005). "Automatic Detection of Features in Ultrasound Images of the Eye", IMTC 2005 – Instrumentation and Measurement Technology Conference, Ottawa, Canada, 17-19. May, 2005.

[2] Xiaoyang, Song; Keou, Song, and Yazhu, Chen "A Computer-based Diagnosis System for Early Glaucoma Screening", Proceedings of 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005.

[3] Rafael, C. Gonzalez, and Richard, E. Woods, Digital Image Processing, 2nd Edition, Pearson Education Asia Publications.

[4] Rafael, C. Gonzalez; Richard, E. Woods, and Steven, L. Eddins, Digital Image Processing using MATLAB®Ⅱ, Pearson Education Asia Publications.

[5] Glaucoma, Research Foundation - Funding Innovative Research to find a cure for Glaucoma.251 Post Street, Suite 600, San Francisco, CA.

*****

AUTOMATIC BRAKING SYSTEM ASSISTED BY IMAGE & VIDEO PROCESSING FOR PEDESTRIAN AND VEHICLE DETECTION USING MATLAB

R. Sakthivel Murugan[100] N. Bose[101]

ABSTRACT

During the next decade, advanced driver assistance systems will play a key role in the challenge of increasing automobile active safety. In this paper, an automatic braking system has been developed in automobiles. When it is fully implemented in vehicles, it could prevent some of the numerous vehicle-pedestrian and vehicle-vehicle accidents.

The system is vision-based and the problem has been to design by vision algorithms by Image and Video Processing which are robust enough to reliably detect any pedestrians and vehicles in a highly cluttered urban environment. Braking is done automatically when the driver gets distracted or when his visibility is poor.

Automatic braking is applied by the output of algorithms using decision logic. The system is also designed for easy future augmentation for, e.g., night driving conditions. This system shows improved robustness with respect to the algorithms that are fast enough to be suitable for on-line operation.

KEYWORDS

Automatic Braking System, Image & Video Processing, Driver Assistance System, Pedestrian Detection, Vehicle Detection etc.

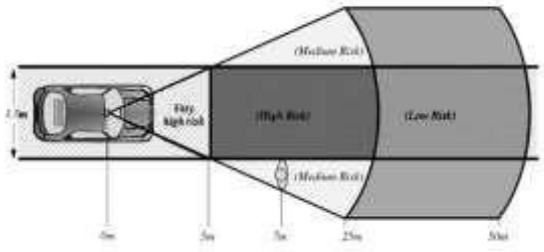
I. INTRODUCTION

The analysis of the statistics of road traffic accidents becomes mandatory when designing Intelligent Transportation System applications. Each year, thousands of pedestrians and cyclists are struck by motor vehicles. To understand the need for a pedestrian detection system, reports of human-vehicle accidents have to be studied. Only in European Union alone about 8.000 pedestrians and cyclists are killed and about 300,000 injured every year. In North America approximately 5,000 pedestrians are killed and 85,000 injured per year. In Japan approximately 3,300 pedestrians and cyclists are killed and 27,000 injured in 2004 [1]. Most of these accidents take place in urban areas where serious or fatal injuries occur at relatively low speed.

The purpose of the paper is to use computer technology, particularly image and video processing, to make for a safer driving experience. The goal of the project is not to create an autonomous vehicle, but to provide input to the driver about road conditions, obstacles and potential hazards, letting the driver retain control of the vehicle while benefiting from the input from the computer system. Such systems are generally referred to as Advanced Driver Assistance Systems (ADAS).
Now, when airbags find a commonplace and passive safety almost has reached its limits, it is time to look toward what can be done to prevent collisions in the first place, especially when it comes to pedestrians. Another approach would be to prevent the actual collision, or at least reduce the energy in a collision either by warning the driver or by the car itself initiating the braking.

In last decade, in addition to the improvements in the road infrastructures (e.g., visibility enhancements, roundabouts, speed controls, better signposting, etc.), a new area of research has received a special focus on ADAS. These are intelligent on-board systems that aim at anticipating and preventing accidents, or at least, minimizing their effects when unavoidable. Examples of ADAS are the Adaptive Cruise Control, which adjusts the own vehicle speed in order to keep a safe gap with the preceding vehicle, or the Lane Departure Warning, which warns the driver in case that the vehicle leaves the lane inadvertently.

Figure-1: The Different Areas of Risk When Driving



Sources: Authors Compilation.

Fig.1 illustrates the typical risky areas to be tackled by Automatic braking system. High risk area, in red, corresponds to a big danger of collision with pedestrians, always depending on the speed of the vehicle. Pedestrians in the medium risk area, in yellow, are likely to cross the front road, so typically no imminent is expected but the system must be aware of them. The low risk area, in green, contains pedestrians with no danger of imminent collision but that must be detected in advance since they stand in the vehicle's path. In regular conditions, the vehicle stopping distance is about 5 m at 30 km/h, increasing up to 12 m at 50 km/h, thus the systems must intelligently focus their techniques on the danger of detecting a pedestrian in these areas.

All India road data [2], shows that 83.5% of the accidents were due to driver's fault, other contributory factors were: mechanical defects in vehicles (3%), pedestrian fault (2.3%), fault of passenger (2.4%), bad roads (1.1%), bad weather (0.9%) and other factors like cattle, fallen trees etc. (6%). Statistics of 83.5% accidents resulting from drivers fault indicates the need of an intensive intervention at host level. So it's necessary to find some safety system to automobiles in order to prevent such accidents by driver's fault.

[100] Assistant Professor, Kamaraj College of Engineering and Technology, Tamil Nadu, India, r.sakthivelmurugan@gmail.com
[101] Professor, MEPCO Schlenk Engineering College, Tamil Nadu, India, irudayaraj67@gmail.com

The above statistical report shows that pedestrians are killed more by accidents. Most of these accidents take place in urban areas where serious or fatal injuries occur even at relatively low speed (within 35 km/h). In order to reduce the accidents occurring to the pedestrians, this work has been undertaken.

In this paper, a particular type of ADAS, pedestrian protection systems (PPS) is focused. The objective of a PPS is to detect the presence of both stationary and moving people and vehicles in a specific area of interest around the moving host vehicle in order to warn the driver, by certain audio alarms and perform braking actions (evasive actions could be an option if the pedestrian surroundings are sensed).

II. LITERATURE SURVEY

Another example of a pedestrian detection and warning system is the PROTECTOR system, part of a European Union research project by DaimlerChrysler Research [3]. The PROTECTOR system is intended to itself stop or slow the car down in the event of a possible impact. The PROTECTOR system calculates trajectories and warns the driver for all pedestrians entering a zone from 5 meters in front of the vehicle to 25 meters ahead of it, which to the PROTECTOR system is the significant zone. It uses two cameras, each with a 30° angle, and has been tested on real roads in Germany, both with actors and "ordinary" pedestrians, the car being driven at a speed of 30 km/h.

The Markovian model of pedestrian behavior [4] is a statistical model of pedestrian behavior, based on a four discrete states Markov chain. The states are "standing still", "walking", "jogging" and "running", all taken from pedestrian paces. Not all transitions are possible. It is for example not allowed to go from jogging to standing still without having walked, or from walking straight to running without having jogging. In this model of pedestrian behavior the current pedestrian pace is calculated given the former one. This is then used to give a speed vector with a norm and an angle, which in turn follows the present state. The behavior of all pedestrians is calculated individually. The aim of this paper is to develop a system that calculates the pedestrian and vehicle trajectories respectively and through that can predict the risk of impact. It is however not yet a working system but only a proposal.

An initial hypothesis generation procedure has been developed [5], which combines two information cues, texture and the outer contour of the pedestrian. The idea of using the outer contour only, is motivated by the different shapes the inner contour can take, for example women with skirts, people with handbags. Since a pedestrian's legs look different depending on where in the walking cycle the images are taken, 'x' a distance metric is used to measure similarity, i.e., to correlate the actual images with the model. The fusion of texture and contour information is made in a temporal dynamic activation field, wherein the final decision of reliable regions is made. It draws upon biomechanical research to set up a model of the human gait and use it as their primary detection scheme. It is an interesting and different piece of work, however, due to fundamental limitations in the approach, such as time delay, not much have been made in this area lately in pedestrian detection. This approach mainly focuses on pedestrians crossing the street being seen from the side, where the walking pattern is visible. This would, however, not suffice as a safety system on its own, since pedestrians on the side walk could rapidly decide to make a street crossing in front of the vehicle, and thus the system would not be able to respond in time.

An impressive work in pedestrian detection has been made based at the Robotics Institute at CMU [6]. The presented system is aimed at transit buses in urban scenes. Their system is based on neural networks for classification and use stereo to make foreground/background separation. The system first separates the image into sub-areas by the stereo range information, where each depth in the image gets its own bounding box. The bounding boxes are preprocessed by the size ratio of a human. Small areas close to each other with similar disparity values are grouped if their combined size complies with the human ratio. Large areas are searched with a window of human size; if nothing is found the area remain unchanged.

Shape based detection has been implemented for pedestrian detection system [7]. The system is shape based in the sense that wavelets are used for feature representation and support vector machine (SVM) for classification. The features encode local intensity differences vertically, horizontally and diagonally. The wavelets extracts the important features of a pedestrians using, e.g., intensity and color, which reduces the influence of noise and non-relevant details in the signatures.

A brute force method [8] is used to detect pedestrians in a captured image, where they shift a search window over all possible locations. Hence, this method is computationally expensive. As a previous approach the method requires 20 minutes per frame examined, not quite real-time. However, after removing the use of color information, reducing the amount of wavelet features and adding stereo to locate regions of interest, their system operates at 10 Hz. It is to be expected that the addition of such information would yield better performance, but of course would slow down the system. The work of the author contains several additions to the field of pedestrian detection. The use of SVM as classifier has been adopted by many groups.

A shape-based pedestrian detection system has been developed, in working with the Urban Traffic Assistant (UTA) system [9]. The core algorithm is based on template matching on the gradient image, where templates are build-up in a hierarchy with general templates near the root and more specific templates at leaf level. The matching is made using a technique called the distance transform, which is applied to an edge detected version of the image. The distance transform gives a gray scale image, where brighter areas are further away from an edge. The template is matched by a threshold on the distance transformed image. This gives a more forgiving matching, than if the template would have been matched directly against the edge image. Following the hierarchy matching is a verification step where the positively matched candidates are run through a neural network, based on an intensity feature representation, in contrast to the templates in the first step. This step is designed to reduce the amount of false matches, but it is not specified how much this step actually affects the final performance of the system. Tracking also is added in later versions of this system. The algorithm is a simplified implementation of a Kalman filter. One of the fundamental problems with this template approach is that it cannot detect occluded pedestrians.

A new method for pedestrian detection was presented [10] and tracking using a single night-vision video camera installed on the vehicle. To deal with the non-rigid nature of human

appearance on the road, a two-step detection/tracking method is proposed. The detection phase is performed by a support vector machine (SVM) with size-normalized pedestrian candidates and the tracking phase is a combination of Kalman filter prediction and mean shift tracking. The detection phase is further strengthened by information obtained by a road-detection module that provides key information for pedestrian validation. Experimental comparisons (e.g., grayscale SVM recognition versus binary SVM recognition and entire-body detection versus upper-body detection) have been carried out to illustrate the feasibility of this approach.

A night vision system has been developed by utilizing infrared technology [11]. A stereo rig with infrared cameras is used to estimate the distance in the scene. The ego-motion of the vehicle is estimated by yaw information from a gyroscope and speed from the vehicle's speedometer. The ego-motion is used to get the relative movement in the scene. Pedestrians are located in the image via a simple mean intensity threshold, which finds areas that are bright (hot) in the image. The areas that passed the threshold are measured in size. Areas ranging in size from head-size to full-body-size are considered to be a pedestrian. The relative movement of these "pedestrians" is compared to the movement of the vehicle to see if they are on a collision course. Only collision course pedestrians are considered. Also it focuses on a system level approach, where they consider response time reductions and presenting the information to the user, not just the detection of pedestrians.

A system has been developed that uses one low resolution infrared camera. [12] An interesting conclusion of this work is that the authors find that pixel-based representations are viable in the infrared domain, compared to the opposite conclusion made by [13] in the visible spectrum, where pixel-based representations perform poorly compared to feature-based.

Another approach is based on images of low resolution infrared cameras [14]. They utilize the fact that humans from most viewpoints are somewhat symmetrical and use the size and aspect ratio of a full standing human to separate pedestrians from other objects in the scene. Bounding box sizes are used to choose which potential pedestrians to process, small bounding boxes are set equal to low information content and thus neglected. Of course, there will be a lot of false positives with that discrimination only. The problem is solved by the use of a set of ad hoc filters and morphological operators to filter out false positives. They also base their system on an initial camera calibration to know the ground plane that is assumed to hold for all future. This kind of calibration assumption will not be very robust, due to movements of the cameras caused by the vehicle. The approach does not use any temporal information, which probably would have made the system more accurate. However, their main focus seems to be to evaluate the possibilities of an infrared camera system.

The Honda Intelligent Night Vision system [15] has details of this system. It is commercially available in Japan since January 2005, tracks and detects pedestrians in the path of the vehicle and calculates their trajectories and positions. The system monitors wheel speed and yaw rate. It also has a daylight sensor and an ambient temperature sensor. The camera monitors the road within a 12° arc and provides auditory and visual alerts using a head-up display. The system measures distance by stereo vision and shape recognition and are capable of providing alerts 30 to 80 meters ahead of the vehicle.

The Intelligent Night Vision system provides an alert when a pedestrian is in vehicle's path or approaching the vehicle's path. It also gives an alert for pedestrians walking along the side of the road, even if the vehicle travels close to the centre line. The vehicle's path is represented by the car's breadth within the 30-80 m area (Honda's patent, 2004). The system on the Honda website only showed detecting pedestrians on country roads, with one or few pedestrians at a time on the road. Honda has developed the world's first Intelligent Night Vision System which could dramatically cut night time pedestrian casualties.

The system is already available in the Japanese market Honda Legend. The new system uses "far infrared" cameras to detect pedestrians in or approaching the vehicle's path and provides the driver with visual and audio cautions to help prevent accidents involving pedestrians. The Intelligent Night Vision System uses images obtained from two far infrared cameras positioned in the lower section of the front bumper to detect the position and movement of infrared heat-emitting objects and determine whether they are in or approaching the vehicle's path. Based on size and shape, the system also determines if the detected object is a pedestrian.

In addition to the conventional night vision function of giving the driver an enhanced view of the road ahead, the system is the world's first to provide cautions that inform the driver of the presence of pedestrians who are on the road or about to cross the vehicle's path. The camera obtains a visual image based on the "far infrared" radiation emitted by humans and other objects. Because it uses far infrared radiation, it is capable of obtaining a viable image without the use of a light source, as is required by visible-light or "near infrared" cameras. The image is then reflected in a mirror positioned on top of the dashboard, which retracts into the dashboard for daytime stowage. The ECU determines pedestrian position and motion based on the image from the cameras, along with vehicle speed and other vehicle information. The system detects pedestrians in or approaching the vehicle's path, and provides caution to the driver via a visual enhancement frame around the pedestrian image and an audio caution.

Mobileye's pedestrian detection technology is based on EyeQ2 systems. Mobileye's [16] has a unique approach to Pedestrian detection lies in the use of monocular cameras only, using advanced pattern recognition and classifiers with image processing and optic flow analysis. Both static and moving pedestrians can be detected to a range of around 30m using VGA resolution imagers. As higher resolution imagers become available range will scale with imager resolution, making detection ranges of up to 60m feasible. Mobileye's first start of production for Pedestrian detection systems is in 2009 on a range of Industrial Powered Vehicles.

In this application EyeQ2 based monocular cameras provided a 360degree all-round Pedestrian Detection system to a range of 15m and will warn the vehicle operator via Audio/Visual warnings of pedestrian in the vehicles path.

In late 2009 Mobileye migrate Pedestrian Detection warning functions to the next generation of consumer product line systems. Mobileye is doing research on full emergency braking for Collision Mitigation for pedestrians where Vision is the key technology for detecting pedestrians. This system is part of the next generation Volvo radar-vision fusion system which also provides Lane Departure Warning and Vehicle Detection with

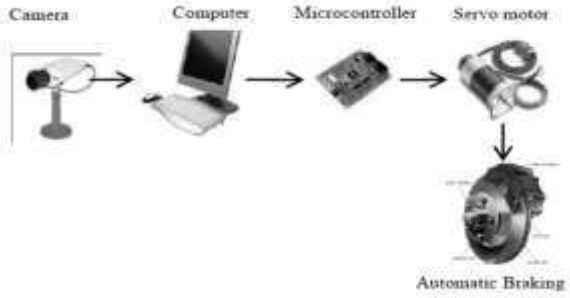
radar-vision fusion for an enhanced Collision mitigation by Braking System on the next Volvo S60 vehicle.

III. METHODOLOGY

This paper mainly focuses on automobile active safety system. A camera will be mounted in front of the automobile prototype model. The vehicle prototype used for this paper is a remote car. The camera detects the moving objects in front of vehicle prototype like pedestrians, moving car etc.

An ultrasonic sensor is mounted on the prototype which is used to measure the distance of the detected or tracked object to the prototype model. The detection and tracking of the moving object by the camera is done by using Image and Video processing techniques. These techniques are done by creating new detection and tracking algorithms in MATLAB software. Using MATLAB Simulation, the results are then integrated with computer and braking is applied automatically to the prototype. The Pre-Crash Brake is applied here with the help of fusion of an ultrasonic sensor and a camera unit. A schematic overview of the automatic braking system is shown in the figure 2.

Figure-2: Process Description



Sources: Authors Compilation.

This brief overview is divided into a hardware section and a software section:

A. Hardware

A schematic overview of the hardware is shown in Fig. 2. The Camera head is mounted on the rear view mirror to get a full view of the traffic scene ahead. The Camera is connected to the main computer via a frame grabber board (FG). All general purpose tasks are run on the main computer, such as vision algorithms and various servers (covered in the software section). Then the main computer is interfaced with micro controller. Based upon the output of the main computer, input is given to the microcontroller. The input of the micro controller may be 0 or 1. Based upon this input the microcontroller actuates the servo motor. The servo motor actuates braking.

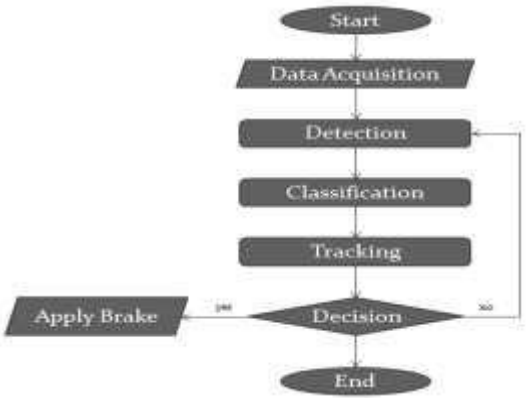
B. Software

The software used for this project for image and video processing section is done by MATLAB software [17]. The toolbox Simulink is used for detection step and also Matlab codings are used for further steps of algorithm. Microcontroller codings should be done by using C-program.

IV. ALGORITHM ARCHITECTURE

The Automatic Braking System constitutes the four main components of algorithms. The detection component is for speed, the classification component for precision, the tracking component for continuity and the decision step for integration. A schematic algorithm of hierarchy of the Automatic Braking System is shown in Fig. 3.

Figure-3: Algorithm Flowchart



Sources: Authors Compilation.

The detection component is used to focus attention on objects that differ from the background based on certain characteristics. Attention characteristics are, e.g. motion, color, depth, and heat differences to the background. The regions which are found by the attention mechanism are segmented, tagged with a probability and sent to the framework.

The framework provides the probability tagged candidates to the classification component, where they are classified. The classification returns a second probability set for the candidates back to the framework. The classification operates at 10–20 Hz depending on workload. Next, the framework requests a third set of probabilities from the tracking component of currently tracked pedestrians. The tracking operates at 60 Hz. The framework fuses the probabilities from the different components to make a final decision of which candidates most probably are pedestrians or vehicles .It give input to audio alerts and also give input whether to apply brake or not.

C. Detection Algorithm

Detection is made by an attention mechanism that focuses attention on regions in an image that adhere to a certain characteristic. The reason for employing an attention mechanism is to reduce the search space in the image. In most cases, the total area in the image covered by interesting objects is fairly small compared to the entire area of the image. Obviously, a reduction of the search area in an image will drastically reduce the computational cost required to find the object, which is important in a real-time system. An ideal attention mechanism captures everything that exhibits the specific characteristics the attention is based on and rejects any areas that do not have those characteristics. E.g., focus of attention by depth ought to find everything that "stands out" from the background, and reject areas that have the same depth as the background. Good detection rate may be at the cost of false positives, which fortunately can be discarded in a following classification step. The extracted candidate regions

are also often referred to as regions of interest (ROI). A basic requirement for an attention mechanism is that the extracted region of interest contains many, and ideally all, instances of an interesting object in the image. The system will never be better than the attention mechanism it is relying on. E.g., an attention mechanism where only half the instances are contained within the regions of interest would yield a system which could never perform better than finding half the objects in the image, no matter how well the rest of the system works. In this sense, the performance of the detection is an upper limit for the detection rate of the integrated system. In general, the higher detection rate required, the larger the extracted regions of interest will be. Forcing one attention mechanism to reach the high required detection rates would lead to a poor segmentation, with large segmented areas of the image as a result. The algorithms used for the detection step is achieved by SAD blockset in Matlab Simulink using SAD method [18]. The sum of absolute differences (SAD) method is a popular technique for motion detection in video processing in Matlab. The SAD block finds the similarity between two input images by performing the sum of absolute differences. The greater the similarity between the two matrices, smaller the SAD values that result.

1): Algorithm for Detection Steps

Table1 shows the five steps of algorithm for the detection algorithm in Automatic braking system.

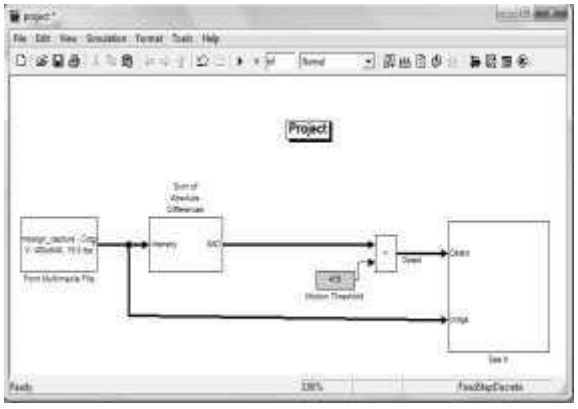
Table-1: Detection Algorithm Sequential Steps

Step 1	Read Video file.
Step 2	Calculate SAD value (Sum of Absolute Difference) of motion
Step 3	If Calculated SAD value of motion > motion threshold (manual), then detect motioned object.
Step 4	The object which is in motion is denoted by red colour in that video.
Step 5	Output the video file in separate colour format.

Sources: Authors Compilation.

2): Simulink Blocksets for Detection Algorithm

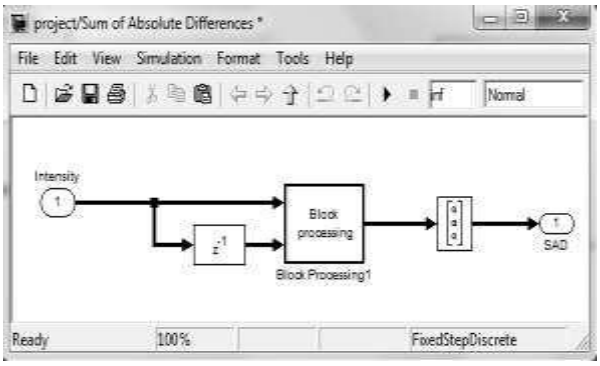
Figure-4: General Layout of Detection Algorithm



Sources: Authors Compilation.

Fig. 4 illustrates the general layout of the detection algorithm. The first blockset is the input of the acquired video signals. Then SAD values are calculated by Sum of absolute difference sub block. A relational operator block is then given which detects the motioned object when SAD value is greater than the Motion threshold value (manual). The detected values are the displayed by a separate output video in last see it block set.

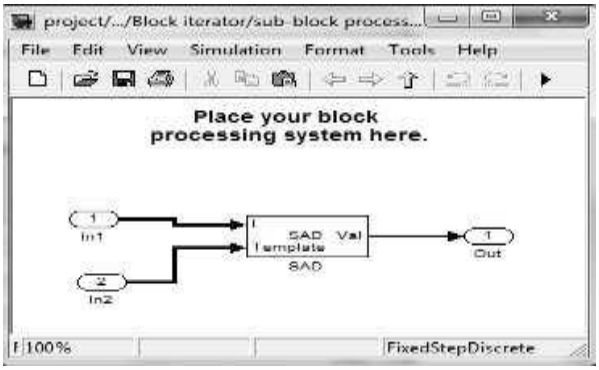
Figure-5: Sub Blocks of SAD Blockset



Sources: Authors Compilation.

Fig. 5 shows the group of blocks inside the Sum of absolute difference blocks. Here a delay mask is inserted and is block processed in a logical way to get the SAD value.

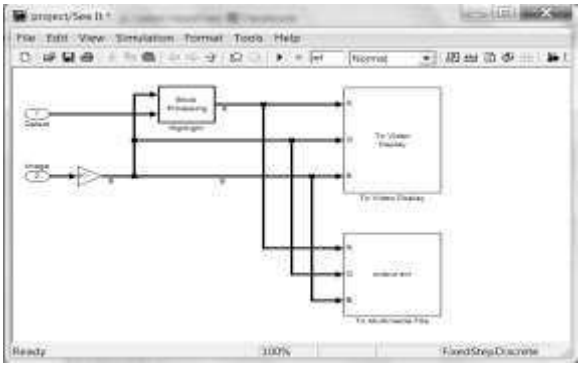
Figure-6: Blockprocessing of SAD Blockset



Sources: Authors Compilation.

Figure-6 shows the block processing in the Sum of Difference sub block. The SAD block finds the similarity between two input images by performing the sum of absolute differences. The greater the similarity between the two matrices, the smaller the SAD values that result the output.

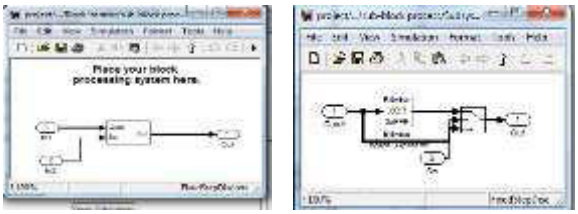
Figure-7: Sub blocks of Output Display (see it) Blockset



Sources: Authors Compilation.

Fig.7 illustrates how the detected motioned objects are to be displayed in a separate video output. A gain operator is used as element wise gain. To display blocks are used to display the output signals with separate RGB components.

Figure-8: (a) Block Processing 2,
Figure-8: (b) Sub Block of Block Processing 2



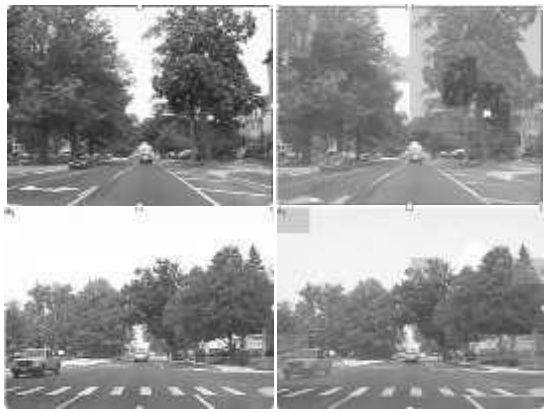
(a) (b)

Sources: Authors Compilation.

Figure 8(a) and figure 8(b) shows the Block processing 2 layout and its sub block in the Output sub block see it. A Bitwise logical operator is used to reduce the fluctuation in the detected output signals. To acquire this a three way switch block is used to get a single output signal.

3): Detection Algorithm Results

Figure-9: Blockprocessing of SAD Blockset



(a) (b)

Sources: Authors Compilation.

Fig. 9 shows the output video frames get by the detection algorithm. Fig. 9(a) shows the original input frame of the acquired video signals and Fig. 9(b) shows the corresponding output frame of video processed detected output video signals.

4): Summary of Detection Algorithm

This chapter has presented the first part of the algorithm 'Detection'. Here the objects which are in motion are detected. This is done by SAD method using the SAD blockset in MATLAB Simulink Library followed by the above sequential connected blocksets.

Here the pedestrian, vehicles and other objects including lamp posts & trees are detected. They resembles in red color in the output video. It means that the detected object is shown in red color. This concludes the description of the detection component.

D. Classification Algorithm

The classification component works on the regions of interest provided by the detection step. The aim of this algorithm is to filter unwanted noises in the output of the detection algorithm. The classification algorithm is achieved by Matlab codes [19]. This stage reduces the analysis area for decision algorithm and thus it improves the decision output time of response. The classification algorithm is achieved by four sequential stages of image processing. Matlab codes should be used for image processing.

1) First Stage

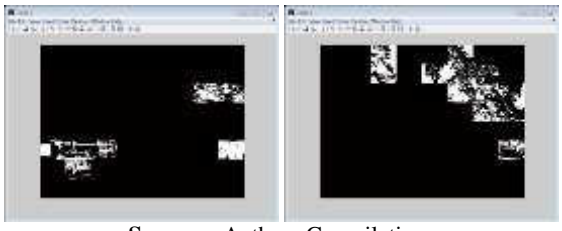
Figure-10: Classification Algorithm-First Stage

Sources: Authors Compilation.

In the first step, the output of the detection algorithm is converted into gray scale images. The red colour ROI of the detection algorithm is having some pixel intensity variation after it gets converted to gray scale. The output of the first step of classification stage is shown in the Fig. 10.

2) Second Stage

Figure-11: Classification Algorithm-Second Stage

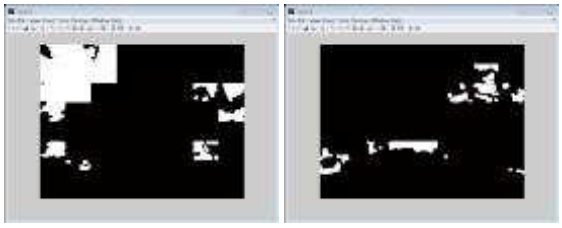


Sources: Authors Compilation.

In the second step, gray scale absolute image subtraction method is implemented between original captured gray scale frame and the detected gray scale frame. So by using certain image processing codes, the red colour ROI of the detection step is get extracted separately in black and white pixels. The output of this step is shown in the Fig. 11.

3) Third Stage

Figure-12: Classification Algorithm-Third stage

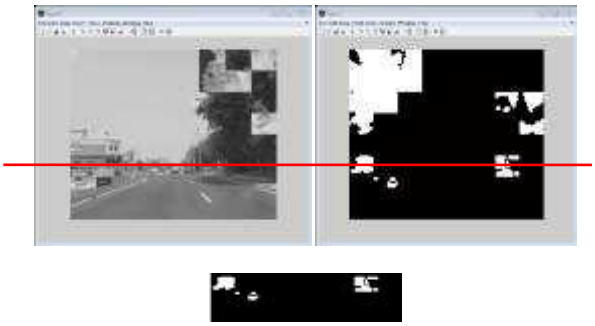


Sources: Authors Compilation.

The output of the second step of classification algorithm is having some mixtures of black and white pixels in the extracted ROI from detected ROI. So in order to go for further process we need to fill the gap between the pixels. Thus the frame contains only a group set of black and white pixels. The output of this stage is shown in the Fig. 12.

4) Final Stage

Figure-13: Classification Algorithm-Last Stage



Sources: Authors Compilation.

In the last stage of classification algorithm, cropping is done as reference to the road plane [20]. This filters the noise of the detected ROI. It filters the detected moving objects like trees, lamppost etc. which are not needed above the road plane. From the cropped area output further more analysis is to be carrying on which increases response time for decision outputs. The output of the last stage of classification algorithm is shown in the Fig.13.

5) Summary of Classification Algorithm

This long chapter has introduced the second part of the automatic braking system algorithm, namely the classification component. The chapter has dealt with four stages of classification algorithm which filters the noises and clutters in the viewing environment. This algorithm is achieved by Matlab codes.

E. Tracking Algorithm

The tracking component is the third stage of automatic braking system algorithm, which is employed to provide continuity to the system. This is concluded by an application of the algorithm to more easily compile ground truth data. The tracking is used to provide continuous processing of the system. The detection component might not always provide a segment for a specific instance of a pedestrian in every frame, and the classification component may not always correctly classify that instance. An additional tracking component, however, balances this with its strong continuity.

F. Decision Algorithm

Decision problems are problems where the answer is either yes (Y) or no (N). Such problems are central to computer and ubiquitous in the world. A decision algorithm is an algorithm that determines the answer to such a problem. A complete decision algorithm is a decision algorithm that always gives the answer in finite time. An incomplete decision algorithm never finishes if the answer is N and May or may not finish if the answer is Y. So, if such an algorithm finishes, the answer is Y. Incomplete algorithms are important because they can often solve significantly larger problem instances than complete algorithms. In automatic braking system algorithm architecture there are two types of decision to be taken.

1) For Audio Alerts

It gives alert signals in the way of Audio alarm, when any moving objects come under risk areas in front of the vehicles. Objects detected in far area are having low intensity area. When it gets nearer to the vehicle then its intensity area increases due to its closeness to the vehicle.

Therefore this intensity area of the detected objects is directly proportional to the distance of that object to the vehicle. The decision for the decision algorithm depends on intensity area in the last stage of classification algorithm. According to it if, the intensity area is greater than 500 pixels in that ROI then it gives alert signals in the form of audio alarms else it won't give any output. It's shown in the Table2.

Table-2: Corresponding Intensity Ares & Its Out for Audio Alerts

Intensity Area	Output	Application
Greater than 500 pixels	Alert Signal	Audio Alarm
Less than 500 pixels	No signal	Nil

Sources: Authors Compilation.

2) For Braking

Figure-14: Cropped Area for Braking Decision

It gives decision logic either to apply brake or not. To achieve this further more cropping is implemented in the predefined cropped area as shown in the Fig. 14.

Sources: Authors Compilation.

For braking, if the intensity area in the newly cropped area is greater than 800 pixels then it gives signals for braking. It shown in the Table3.

Table-3: Corresponding Intensity Ares and Its Out for Braking

Intensity Area	**Output**	**Application**
Greater than 800 pixels	Brake signal	Braking
Between 800~500 pixels	Alert signal	Audio Alarms
Less than 500 pixels	No signal	Nil

Sources: Authors Compilation.

3) Summary of Decision Algorithm

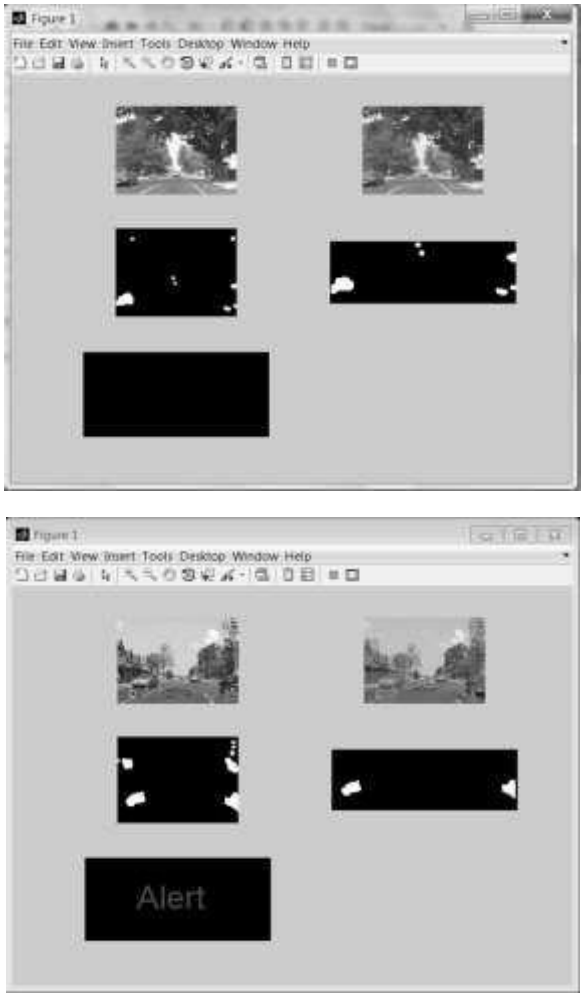
This algorithm has presented the novel conceptual work with decision logic with parallel information flow. The developed decision is applied to the automatic braking system and the communication needed between components has been discussed. If the system won't give output for decision then it will go to detection step as shown in figure 3.

V. RESULTS AND DISCUSSION

The output of the software architecture for the automatic braking system using MATLAB is shown in the figure 15. In this the MATLAB window consists of five output modules. The first module shows the original data acquisitioned live video captured by the camera.

Second module shows the detected ROI output by detection algorithms. Third module shows the output of third stage of classification algorithm. The fourth module shows the output last stage of classification algorithm. The last module just displays the output of the decision for the automatic braking system.

Figure-15: MATLAB Output Module



Sources: Authors Compilation.

CONCLUSIONS

Intelligent vehicles represent a key technology for reducing the number of accidents between pedestrians and vehicles. This automatic braking system is particularly valuable when the driver is distracted or visibility is poor.

In this work, automatic braking is achieved by Image and video processing. The paper mainly focuses on the algorithm for automatic braking. The condition in which this study was made was sufficient for its aim.

In future studies however it would be advisable to add more factors such as darkness or bends in the road. It could also be an idea to allow participants to drive for themselves and then record the actual behaviour, such as turning the wheel, slowing down and moving the foot to the brake pedal.

VI. REFERENCES

[1] Pedestrian Safety Global Technical Regulation Preamble, 2004, In Transport Division, World Forum for Harmonization of Vehicle, October 2004. www.unece.org/trans/doc/2005/wp29grsp/ps-170e.pdf.

[2] M., D. Arvind Kumar, (2008). "Fatal Road Traffic Accidents and their Relationship with Head Injuries", Indian Journal of Neurotrauma (IJNT), Volume: 5, Number: 2, pp. 63-67, 2008.

[3] D., M. Gavrila; J., Giebel, and S., Munder, (2004). "Vision-Based Pedestrian Detection: The PROTECTOR System", IEEE Intelligent Vehicles Symposium, Italy, 2004.

[4] S., Wakim, and Capperon, J. Oksman, (2004). "A Markovian Model of Pedestrian Behaviour", IEEE International Conference on Systems, Man and Cybernetics, The Hague, Netherlands, 2004.

[5] C., Curio; J., Edelbrunner; T., Kalinke; C., Tzomakas, and W., Seelen, (2000). "Walking Pedestrian Recognition", IEEE Transactions on Intelligent Transport Systems, Volume: 1:3, October.

[6] L., Zhao, and C., E. Thorpe, (2000). "Stereo-and Neural Network-based Pedestrian Detection", IEEE Transaction on Intelligent Transportation Systems, Volume: 1, Number: 3, pp. 148 – 154, September.

[7] M., Oren; C., Papageorgiou; P., Sinha; E., Osuna, and T., Poggio, (1997). "Pedestrian Detection using Wavelet Templates", Proceedings of Conference on Computer Vision & Pattern Recognition, pp. 193-199.

[8] C., Papageorgiou; Evgeniou, X, and T., Poggio, (1998). "A Trainable Pedestrian Detection System", IEEE Intelligent Vehicles Symposium, pp. 241–246, October.

[9] D., M. Gavrila, (2000). "Detection from a Moving Vehicle", European Conference on Computer Vision (ECCV), Volume: 2, pp. 37– 49.

[10] Fengliang, Xu; Xia, Liu, and Kikuo, Fujimura, (2005). "Pedestrian Detection and Tracking with Night Vision", IEEE Transactions on Intelligent Transportation Systems, Volume: 6, Number: 1, March.

[11] T., Tsuji; H., Hattori; M., Watanabe, and N., Nagaoka, (2002). "Development of Night Vision System", IEEE Transactions on Intelligent Transportation Systems, Volume: 3, Issue: 3, September.

[12] H., Nanda, and L., Davis, (2002). "Probabilistic Template Based Pedestrian Detection In Infrared Videos", IEEE Intelligent Vehicle Symposium, June.

[13] C., Papageorgiou, and T., Poggio, (1999). "Trainable Pedestrian Detection", IEEE International Conference on Image Processing, Volume 4, pp. 35–39.

[14] M., Bertozzi; A., Broggi; P., Grisleri; T., Graf, and M., Meinecke, (2003). "Pedestrian Detection in Infrared Images", IEEE Intelligent Vehicles Symposium, pp. 662 – 667.

[15] The Honda Intelligent Night Vision System, (2004), Top Gear News, http://www.topgear.co.uk/news/2006/01/18/honda-night-vision-system/

[16] Mobile Eye's Pedestrian Detection, Netherland, September, (2010) http://www.mobileye.com/manufacturerproducts/applications/forward/pedestrian-protection

[17] Matlab 2010b - Software Manual.

[18] Matlab 2010b-Product Help.

[19] http://www.mathworks.com/matlabcentral/

[20] David, Gerónimo, and Angel, D. Sappa, (2010). "2D–3D-based On-Board Pedestrian Detection System", Computer Vision and Image Understanding, Volume: 114, Issue: 5, pp. 583-595, February.

*****

A SIMPLE BOUNDARY EXTRACTION TECHNIQUE FOR IRREGULAR PUPIL LOCALIZATION

Jincy J. Fernandez[102] G. Annapoorani[103] R. Krishnamoorthy[104]

ABSTRACT

A simple boundary extraction technique for irregular pupil boundary localization is proposed in this paper by designing a new approach that extracts irregular pupil boundary points from the edge map using 4-neighborhood connected component analysis. Initially, Canny's edge detection is adopted to extract the edge map in the eye image. Then the proposed boundary extraction technique is employed to extract the pupil boundary points perfectly and it is independent of the specific eye image characteristics such as pupil deformation, poor contrast, poor brightness, etc. Proposed approach is adaptable to all monochrome databases and exhibits encouraging results when compared with the existing techniques.

KEYWORDS

Biometrics, Iris Recognition, Iris Segmentation, Pupil Localization, Boundary Extraction etc.

INTRODUCTION

With an ever-increasing prominence on security, automated personal identification based on biometrics has been receiving widespread attention over the past decade. Biometric methods, which identify people based on physical or behavioral characteristics, are of interest because people cannot forget or lose their physical characteristics in the way that they can lose passwords or identity cards. There has been an explosion of interest in iris biometrics in recent years.

It is suggested in [1-3] that human irises are as distinct as fingerprint for different individuals, leading to the idea that iris patterns contain unique identification features. The minute details of the iris patterns are believed to be determined randomly during the fetal development of the eye. Iris recognition has great advantages, such as variability, stability and security. The purpose is real-time with high confidence recognition of a person's identity by mathematical analysis of the random patterns that are visible within the iris of an eye from some distance. Iris recognition comprises of three basic stages: iris acquisition, preprocessing and recognition of the iris region. Preprocessing includes pupil localization, sclera localization, normalization and enhancement.

The pupil localization is a critical stage to the success of any iris recognition system, since pupil data that is falsely represented as iris pattern will corrupt the biometric templates generated, resulting in poor recognition rates. Since the existing pupil localization methods are unable to obtain valuable information from an eye image and pupil localization plays significant role in iris recognition, a new irregular pupil localization technique is presented in this paper. Initially an edge map computation is designed with Canny's edge detection. Also, a new and simple boundary extraction approach that finds the pupil boundary points accurately, with the help of labeling the components is proposed. This paper is organized as follows: In Section 2, we present the literature review. In Section 3, we present edge map extraction technique. In Section 4, we present the proposed boundary extraction technique. In Section 5, we present the measure for quantitative analysis. In Section 6, we present the experimental results and comparative results of the proposed design strategy with existing technique. In Section 7, we present conclusions.

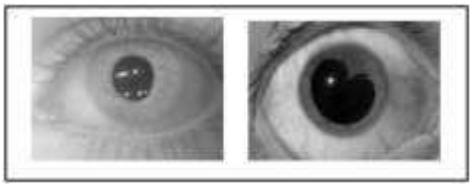
LITERATURE RVIEW

In general there are six fundamental approaches for pupil localization. They are (i) Template-Based Localization, (ii) Circular Boundary Localization, (iii) Cluster Based Localization, (iv) Elliptical Boundary Localization, (v) Combination of Boundary Based Approach and Template Based Approach, and (vi) Irregular Boundary Localization.

The template-based strategy usually involves the maximization of parametric equations. Early works on template based strategy include integro-differential operator by Daugman [4-7], threshold technique by Zhu et al. [8], bisection method by Lim et al. [9], high threshold technique by Bhola Ram Meena [10], threshold based Freeman's chain code by Vatsa et al. [11], edge based virtual circle by Boles et al. [12], integro-differential operator by Sanchez-Avila et al. [13-14], threshold based ring mask by Lye Liam et al. [15], active contour model by Ritter et al. [16] and Daugman [17], etc. Existing state of art on template-based strategy requires reconfiguration of parameters per each database. However, it can be observed that it frequently fails when images do not have sufficient intensity separability between iris and sclera regions.

The circular boundary approach [18] begins with the construction of an edge-map followed by the application of circle geometric form fitting algorithm. This is the most common methodology, being proposed with several minor variants as reported in [19-25]. The circular boundary approach depends on specific image characteristics such as brightness, and contrast. It was reported that edge-detector algorithm and the necessary parameters used for tuning it are critical factors for localization accuracy. Also it is assumed that the boundary of the pupil localization is circle. But it can be noticed that the circular assumption of the contours can lead to inappropriate boundary detection, as evident from figure 1.

Figure-1: Illustration of Imperfect Circularly Pupil Boundaries



Sources: Authors Compilation.

[102] Anna University of Technology, Tamil Nadu, India, jincy.fernandez@gmail.com
[103] Assistant Professor, Tamil Nadu, India, pooranikrish@gmail.com
[104] Dean, Department of Computer Science and Engineering; Research Scholar, Anna University of Technology, Tamil Nadu, India, gkrsih26@gmail.com

The third category viz., cluster based localization techniques are a variant of circular boundary based approach. The cluster based localization begins with clustering/ normalization process, and then edge detection and boundary localization are carried out. Typical cluster based localization includes texture segmentation by Cui et al. [26], GLCM pattern analysis technique by Bachoo and Tapamo [27], normalization based Hough transform by Proenca et al. [28], etc. Even though it has better edge detection capability due to the introduction of normalization procedure, its pupil localization is based on the geometric form fitting techniques. Also, the cluster based localization requires more computation time because it includes preprocessing steps like clustering/ normalization procedure, edge detection and specific boundary localization.

To solve a forementioned problem, some researchers assumed that the pupil boundary is elliptical. It begins with the construction of an edge-map followed by the application of elliptical geometric form fitting algorithm. This is the most recent methodology, being proposed with several minor variants as reported in [29-34].

However, the elliptical assumption is also flawed. In the non-ideal iris images, which are off-angle, motions blur and noisy, pupils appear to be non-elliptical. And even in the ideal iris images, some shapes of pupils are not perfect ellipses. It has been observed that the elliptical assumption of the contours can lead to inappropriate boundary detection in the case of non-ideal and few ideal iris images. There are some literatures that used the combination of boundary based approach and template based approach [35].

In recent years only, it is well recognized that the pupil boundary is irregular and very few works are reported in the literature. Rakshit et al. [36] devised a method of modeling non-circular pupil boundaries using a Fourier-based shape description. Huang et al. [37] proposed irregular pupil boundary localization using radial-suppression edge detection, threshold, and self-adaptive boundary localization.

However, these irregular pupil boundary localization techniques require more computation time because it includes many preprocessing steps like best fit search, selection of wavelet functions, calculation of the modulus of the wavelet transform, radial non-maxima suppression, and edge threshold.

Although the aforementioned approaches follow different approaches, all these methods assume pupil as some curvature form. However, during observation, it is found that the pupil boundaries in the iris images are not always follow specific shape. Hence, in this paper, a new and simple pupil localization scheme is proposed that detects the pupil boundary adaptively regardless of the shapes of the boundary by extracting an edge map and the same are presented in the next section.

EDGE MAP EXTRACTION

In order to effectively extract a pupil boundary, it is essential to define the contour characteristics. In general, a pupil boundary is a closed and continuous curve, but irregular in shape. In this work, Canny's edge detector is adopted as reported in [38]. It starts by a blur operation followed by the construction of a gradient map for each image pixel.

A non-maximal suppression stage sets the value of 0 to all the pixels of the gradient map that have neighbors with higher gradient values. Further, the hysteresis process uses two defined values to classify some pixels as edge or non-edge. Finally, edges are recursively extended to those pixels that are neighbors of other edges and with gradient amplitude higher than a lower threshold.

PROPOSED BOUNDARY EXTRACTION

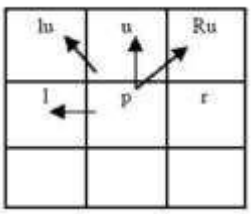
In this proposed work, the boundary extraction scheme is designed in such a way that it considers the edge points from the edge map. The basic assumption of the proposed work is that a useful number of points can be found on the pupil boundary, which are not necessarily equally spaced. Even though edge detection enables proper estimation of the edge points on the pupil boundary, it is required to finely localize the pupil boundary for the purpose of iris segmentation.

A connected component is a set of connected pixels that share a specific property, *V*. Two pixels, *p* and *q,* are connected if there is a path from *p* to *q* of pixels with property *V*. A path is an ordered sequence of pixels such that any two adjacent pixels in the sequence are neighbors. It is possible to find a path between any two pixels within any of the objects, but not between pixels in different objects.

We assume that all of the pixels have a value of the property *V* and that we are interested in those pixels with nonzero values. Those with value zero are assumed to be the image background for each pixel, p. If the pixel value is 1, then check whether its four neighbors are having same label. If so, same label is assigned to pixel *p.*

Each pixel in the image has been labeled with values that correspond to the property of interest. Constructing a connected component consists of growing sets of pixels that are connected and have the and so on. This leads to a random stepping through the image, and is somewhat inefficient. Scanning the image in a specified order can develop a more systematic and efficient technique.

Figure-2: Connected Neighbourhood Analysis



Sources: Authors Compilation.

a) Image Scanning and Labeling*:* Let us first introduce some notation. The image will be represented by an Array *A* that has *N* columns and *M* rows. $A[x][y]$ refers to the element in column y and row x, with $x \square \{0, M-1\}, y \square \{0, N-1\}$. We assume that each pixel has a value.

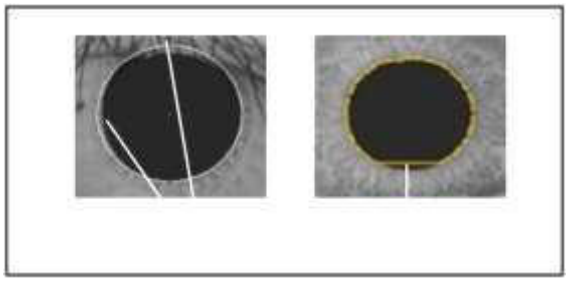
Let Q be an array that is the same size as A. We use this array to hold labels of the connected components in the image. Let *L* be the label index. We start with *L* 0 and increment *L* whenever we want to create a new connected component label. The goal is to end up with all of the pixels in each connected component having the same label and all of distinct connected components having different labels. While scanning the first pixel in image,

check pixel value of $A[0][0]$. If it is nonzero, then increment label index and update this label details in $Q[0][0]$. While scanning the first row, check the pixel value of $A[0][y]$ is nonzero and equals to left neighbor $A[0][y-1]$, assign the label of left neigbour value and update this label details in $Q[0][y-1]$; otherwise increment label and update this label details in $Q[0][y-1]$. While scanning the first column, check the pixel value of $A[x][0]$ is nonzero and equals to $A[x-1][0]$, assign the label of top neigbour value and update this label details in $Q[x][0]$; otherwise increment label and update this label details in $Q[x][0]$. While scanning the remaining pixels other than first row and first column, analyze the left l, top u, left upper diagonal ul and right upper diagonal ur neighborhood pixel values as mentioned in Figure 2. If the pixel value of $A[x][y]$ is nonzero, $A[][xy]!\ l, A[][xy]!\ u, A[][xy]!\ ul$, and $A[x][y]!\ ur$, then increment label index and update label details in $Q[x][y]$. If the pixel value of $A[x][y]$ is nonzero and $A[x][y]\ \ l$, then set $Q[x][y]\ Q[x][y-1]$. If the pixel value of $A[x][y]$ is nonzero, $A[x][y]!\ l$ and $A[x-1][y]\ \ l$, then set $Q[x][y]\ Q[x-1][y]$. If the pixel value of $A[x][y]$ is nonzero, $A[x][y]!\ l$ and $A[x-1][y-1]\ \ 1$, then set $Q[x][y]\ Q[x-1][y-1]$. If the pixel value of $A[x][y]$ is nonzero, $A[x][y]!\ l$ and $A[x-1][y+1]\ \ l$, then set $Q[x][y]\ Q[x-1][y+1]$.

In the course of image scanning and labeling, it is possible that some of the pixels in the same connected component will end up with different labels. These different label values will be identified and resolved at the end of the labeling process. If the neighbors are having same values but different labels, then find the minimum label, *min* and replace the large label with *min*. The labeled connected component with the highest count is marked as irregular pupil boundary points.

same value of a property. This could be accomplished by first finding a pixel with a given property value, then looking at all its neighbors, labeling each that has the same value as being in the same component,

Figure-3: False-Covered Area



Sources: Authors Compilation.

PERFORMANCE MEASURE

The performance of the proposed pupil localization technique is reported by computing the false-covered area, as reported in [37]. As shown in figure 3, False-covered area is computed by adding over-covered area and uncovered area, which is defined as $FC\ OC \sqcap UC$ … (1) where *OC* is the number of extra pixels other than the pupil area inside the pupil boundary and *UC* is the number of pupil pixels outside the pupil boundary.

EXPERIMENTS AND RESULTS

The proposed boundary extraction technique for pupil localization has been experimented with images from the standard iris image databases [39]. Three sample images viz. 001_1_1, 018_2_3 and 107_1_3 images, which are of (320X280) with pixel values in the range 0-255 are shown in figure 4(a), 4(b), and 4(c). The input images are subjected to the Canny's edge detection process. The corresponding resultant edge maps for test images shown in figure 4, after edge map construction process with thresholding the magnitude of the image intensity gradient, are shown in figure 5a), 5(b), and 5(c). The derived edge map is subjected to the proposed boundary extraction technique, as described in section 4.

The edge points in the edge map are considered for boundary extraction, and neighborhood points of each edge point are is evident from the experimental results that only fractions of search space are evaluated instead of an exhaustive search space, which greatly accelerates the search procedures while guaranteeing pupil boundary points.

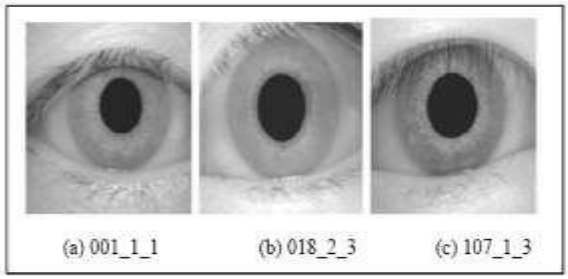
The corresponding boundary extraction results for the test images shown in figure 4, after boundary extraction process with proposed approach are shown in figure 6(a), 6(b), and 6(c) respectively. The corresponding pupil localization results for the test images shown in figure 4, after fusion of boundary extraction results and test images are shown in figure 7(a), 7(b), and 7(c) respectively.

In order to measure the efficiency of the proposed pupil localization, we have conducted experiments with R.P Wildes's methodology [18] and Huang's methodology [37]. The corresponding pupil localization results for the test images shown in figure 4, after fusion of boundary extraction results and test images are shown in figure 8 and 9 respectively. It is analyzed and labeled as described in section 4.1.

The irregular pupil is identified by identifying maximal connected component among connected components. It observed from the figure 8 that the method of Wildes cannot localize the pupil accurately. It is observed from the figure 9 that the method of Huang can localize the pupil accurately. In order to evaluate the performance of the proposed pupil localization algorithm with different localization methods [18] [37], 100 iris images in CASIA database are employed to test the accuracy of pupil localization.

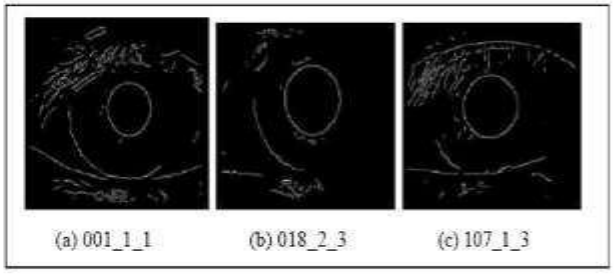
To quantitatively evaluate the pupil localization, we evaluated the pupil localization results using false-cover measure. In Table 1, we give the percentage of covered pupil area in terms of over-covered area, uncovered area and false-covered area. Table 1 shows the visual effect of results of different pupil localization methods. It can be observed from Table 1, the proposed technique is found to be better than the Wilde's methodology and is comparable with Huang's methodology.

Figure-4: Test Images



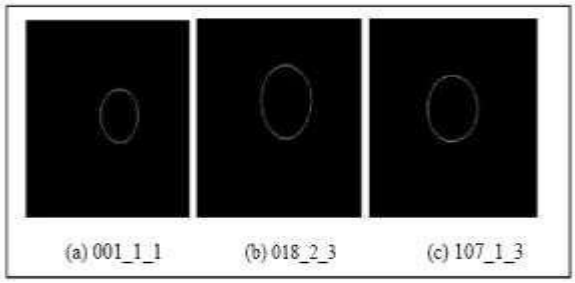
Sources: Authors Compilation.

Figure-5: Edge map with Canny's Edge Detector



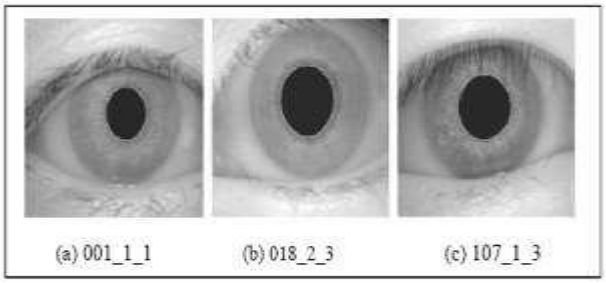
Sources: Authors Compilation.

Figure-6: Boundary Extraction Results with connected Component Analysis



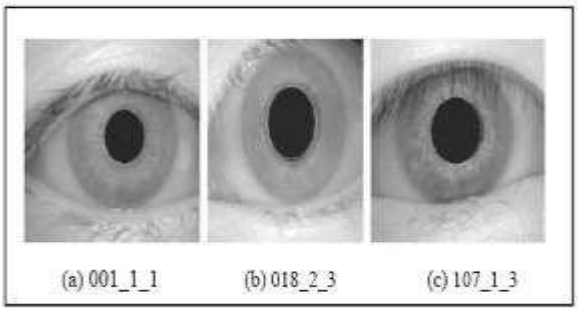
Sources: Authors Compilation.

Figure-7: Pupil Localization Results With Proposed Scheme



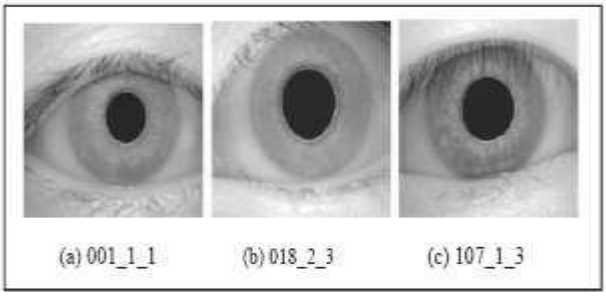
Sources: Authors Compilation.

Figure-8: Pupil Localization Results with Wilde's Methodology



Sources: Authors Compilation.

Figure-9: Pupil Localization Results with Huang's Methodology



Sources: Authors Compilation.

Table-1: Covered Pupil Area with Different Methods of Pupil Localization

Methods	Over-covered area OC (%)			Uncovered area UC (%)			False-covered area		FC (%)
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Proposed Method	0.13	0.11	0.15	0.10	0.11	0.12	0.12	0.16	0.26
Wildes's Method	2.56	0.44	7.52	2.81	0.55	6.25	3.75	0.88	9.01
Huang's Method	0.14	0.12	0.22	0.17	0.16	0.22	0.22	0.17	0.27

Sources: Authors Analysis.

CONCLUSIONS

In this paper, we have proposed a new boundary extraction technique for localizing the pupil boundary from edge map of eye images. Canny edge detection has been used for computing binary edge map.

Based on the Canny's binary edge map, a boundary extraction method is proposed, which detects the pupil boundary adaptively regardless of the shapes of the boundaries. It alleviates the computational burden of the conventional boundary extraction techniques while still seeking a highly accurate localization outcome.

REFERENCES

[1] P., Kronfeld, (1968). "The Gross Embryology of the Eye", The Eye, Volume: 1, pp. 1-66, 1968.

[2] L., Flom, and A., Safir, (1987). "Iris Recognition System", U.S.Patent: 4, 641, 349.

[3] H., Davson, (1990). "Davson's Physiology of the Eye", MacMillan, London.

[4] J., Daugman, (1993). "High Confidence Visual Recognition of Persons by a Test of Statistical Independence", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 15, Number: 11, pp.1148-1161.

[5] J., Daugman, (2001). "Iris Recognition", American Scientist Journal, Volume: 89, pp: 326–333, http:// www. americanscientist.org/articles/01articles/daugman.html

[6] J., Daugman, "Demodulation by Complex-Valued Wavelets for Stochastic Pattern Recognition", International Journal on Wavelets, Multiresolution and Information Processing, Volume: 1, Number: 1, pp. 1-17.

[7] J., Daugman, (2004). "How IRIS Recognition Works", IEEE Transactions on Circuits and Systems for Video Technology, Volume: 14, Number: 1, pp. 21-30.

[8] Y., Zhu; T., Tan, and Y., Wang, (2000). "Biometric Personal Identification Based on Iris Patterns", Proceedings of 15th International Conference on Pattern Recognition, Volume: 2, pp. 805 – 808.

[9] S., Lim; K., Lee; O., Byeon, and T., Kim, (2001). "Efficient Iris Recognition through Improvement of Feature Vector and Classifier", Journal of Electronics and Telecommunication Research Institute, Volume: 23, Number: 2, pp. 61–70.

[10] Bhola, Ram Meena, (2004). "Personal Identification based on Iris Patterns", Ph.D Thesis, Department of Computer Science and Engineering, Indian Institute of Technology, Kanpur.

[11] M., Vatsa; R., Singh, and A., Noore, (2005). "Reducing the False Rejection Rate of Iris Recognition using Textural and Topological Features", International Journal of Signal Processing, Volume: 2, Number: 2, pp. 66-72.

[12] W., W. Boles, and B., Boashash, (1998). "A Human Identification Technique using Images of the Iris and Wavelet Transform", IEEE Transactions on Signal Processing, Volume: 46, Number: 4, pp. 1185-1188.

[13] C., Sanchez-Avila, and R., Sanchez-Reillo, (2002). "Multiscale Analysis for Iris Biometrics", Proceedings of 36th International Conference on Security Technology, Volume: 1, pp. 35-38.

[14] Sanchez-Avila, and Sanchez-Reillo, R, (2005). "Two Different Approaches for Iris Recognition using Gabor Filters and Multiscale Zero-Crossing Representation", Pattern Recognition, Volume: 38, pp. 231-240.

[15] Lye, Liam; Ali, Chekima; Liau, Fan, and Jamal, Dargham, (2002). "Iris Recognition Using Self-Organizing Neural Network, Proceedings of IEEE Student Conference on Research and Developing Systems, pp.169-172.

[16] N., Ritter; J., Cooper; R., Owens, and P., P. Van Saarloos, (1999). "Location of the Pupil-Iris Border in Slit-Lamp Images of the Cornea", Proceedings of International Conference on Image Analysis and Processing, pp. 740-745.

[17] J., Daugman, (2007). "New Methods in Iris Recognition", IEEE Transactions on Systems, Man, and Cybernatics – Part B: Cybernatics, Volume: 37, Number: 5, pp. 1167-1175.

[18] Wildes, R. P., (1997). "Iris Recognition: An Emerging Biometric Technology", Proceedings of the IEEE, Volume: 85, pp.1348-1363.

[19] L., Ma; T., Tan; Y., Wang, and D., Zhang, (2003). "Personal Identification based on Iris Texture Analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 25, Number: 12, pp. 1519-1533.

[20] L., Ma; T., Tan; Y., Wang, and D., Zhang, (2003). "Efficient Iris Recognition by Characterizing Key Local Variations", IEEE Transactions on Image Processing, Volume: 13, pp. 739-750.

[21] M., Nabti, and A., Bouridane, (2008). "An Effective and fast Iris Recognition Based on A Combined Multiscale Feature Extraction Technique", Proceedings of Pattern Recognition, Volume: 41, pp. 808 – 879.

[22] D., M. Monro; Soumyadip, Rakshit, and D., Zhang, (2007). "DCT-Based Iris Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 29, Number: 4, pp.586-595.

[23] Libor, Masek, "Recognition of human iris patterns for biometric identification", 2003, http://www.csse.uwa.au/~pk /student projects /libor

[24] Xiaofu, He, and Pengfei, Shi, (2007). "A New Segmentation Approach for Iris Recognition Based on Hand-Held Capture Device", Pattern Recognition, Volume: 40, pp.1326-1333.

*****

EVCS FOR GENERAL ACCESS STRUCTURE USING EMBEDDING

Mary Mareena P. V.[105] S. Shobana[106]

ABSTRACT

A visual cryptography scheme (VCS) is a kind of secret sharing scheme which allows the encoding of a secret image into n shares distributed to n participants. The beauty of such a scheme is that a set of qualified participants is able to recover the secret image without any cryptographic knowledge and computation devices. Conventional visual secret sharing schemes generate noise-like random pixels on shares to hide secret images. It suffers a management problem, because of which dealers cannot visually identify each share. This problem is solved by the extended visual cryptography scheme (EVCS), which adds a meaningful cover image in each share. However, the previous approaches involving the EVCS for general access structures suffer from a pixel expansion problem. In addition, the visual cryptography (VC)-based approach needs a sophisticated codebook design for various schemes. In this paper, we propose a general approach to solve the above mentioned problems; the approach can be used for binary secret images in non computer-aided decryption environments.

The proposed approach consists of two phases. In the first phase, based on a given access structure, we construct meaningless shares using an optimization technique and the construction for conventional VC schemes. In the second phase construction of EVCS which is realized by embedding random shares into meaningful covering shares, and we call it the embedded EVCS. The shares of the proposed scheme are meaningful images, and the stacking of a qualified subset of shares will recover the secret image visually. To improve the visual quality of the share images we develop an algorithm based on the simulated annealing (SA) approach to solve the proposed mathematic optimization formulation for the GAS problem. We transform and relax the original mathematical model for the GAS solver to simplify solution procedures. Our solution approach adopts an iterative improvement framework. First, we treat the decision variable as a given variable and the original optimization problem is, therefore, transformed into a decision problem. In each iteration, we try to find a solution with a given and refine according to the possibility of a solution being found or not. The display quality of the recovered image is very close to that obtained using conventional VC schemes.

KEYWORDS

Wireless Sensor Network, Convergecast, Sink Nodes, Linear Coding, Deployment etc.

INTRODUCTION

Secret sharing was invented by both Adi Shamir and George Blakley independently in 1979.Secret sharing is a way to distributing secret among a group of members such that each member owns a share of the secret. Only some specific combinations of shares can reconstruct the secret. Individual shares reveal nothing about the secret. Visual Secret Sharing is an encryption technique that doesn't need complex calculations in order to decrypt a message. The ciphertext and the key consist of transparencies (or images). When properly stacked these transparencies (or stacked by image editing package), one can reveal the plaintext.

Visual cryptography (VC), which was proposed by Naor and Shamir, allows the encryption of secret informationin the image form. By applying the concept of secret sharing a secret image can be encrypted as different share images printed on transparencies, which are then distributed to participants. By stacking transparencies (shares) directly, the secret images can be revealed and visually recognized by humans without any computational devices and cryptographic knowledge. On the other hand, any one share or a portion of shares canleak nothing related to the secret image. VC is a very good solution for sharing secrets when computers cannot be employed for the decryption process. The (k, n) VSS can be defined as follow:

(1) n secret shares are constructed according to the original image, and the shares are also called shadows.
(2) k is the threshold number of shares required to reconstruct the original image. That is, the original image can be decoded with any combination of k shadows (2-1); yet no information can be captured with any combination of k' shadows, where k' < k (2-2).

Following is an Example of (2, 2) Schemes:

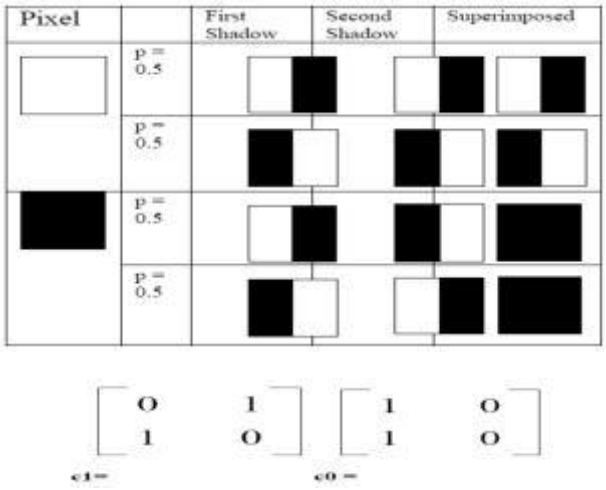
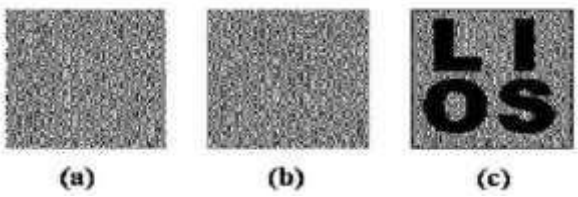


Figure-1: An example of (2, 2) Schemes



Sources: Authors Compilation.
Example of traditional (2, 2) - VCS with image size 128* 128.

[105] PG Scholar, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, pvmareena@gmail.com
[106] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, shobs1812@gmail.com

Conventional VSS schemes generate noise-like random pixels on sharesto hide secret images. In this manner, the secret can be perfectly concealed on the share images. However, these schemes suffer from a management problem—dealerscannot identify each share visually. Hence, researchers have developed the extended visual cryptography scheme (EVCS), also known as the friendly VC scheme, which adds a meaningful cover image on each share to address the management problem.

The pixel expansion problem is a common disadvantage with most of the VSS schemes. When the VC-based approach is employed, each secret pixel within a secret image is encrypted in a block consisting of subpixels in each constituent share image. Thus, the area of a share is times that of the original secret image. The contrast of the recovered images will be decreased to simultaneously. The pixel expansion problem not only affects the practicability of storage / transmission requirements for shares but also decreases the contrast of the recovered secret images.

Generally, an EVCS takes a secret image and original share images as inputs, and outputs shares that satisfy the followingthree conditions: 1) any qualified subset of shares can recoverthe secret image; 2) any forbidden subset of shares cannot obtain any information of the secret image other than the size of the secret image; 3) all the shares are meaningful images. EVCS can also be treated as a technique of steganography. One scenario of the applications of EVCS is to avoid the custom inspections, because the shares of EVCS are meaningful images, hence there are fewer chances for the shares to be suspected and detected.

LITEARTURE SURVEY

[1] Visual Cryptography

In this paper M. Naor and A. Shamir suggested that Visual Cryptography is a special encryption technique to hide information in images in such a way that it can be decrypted by the human vision if the correct key image is used. The technique was proposed by Naor and Shamir in 1994. Visual Cryptography uses two transparent images. One image contains random pixels and the other image contains the secret information. It is impossible to retrieve the secret information from one of the images. Both transparent images and layers are required to reveal the information. The easiest way to implement. Visual Cryptography is to print the two layers onto a transparent sheet. In the overlay animation you can observe the two layers sliding over each other until they are correctly aligned and the hidden information appears.

[2] Constructions and Properties of K-Out-of-N Visual Secret Sharing Schemes

In this paper E. R. Verheul and H. C. A. v. Tilborg introduced the notion of visual *k* out of *n* secret sharing schemes. The idea is that an image (e.g. picture or text) is transformed into *n* transparencies (shares), in such a way that if one puts any *k*-tuple of transparencies on top of each other, the original image is again visible, while with any (*k ¡* 1)-tuple of transparencies no information about the original image is released (in the sense that any possibility is equally likely).

Sharing schemes are useful in military and civilian applications. In the traditional Secret Sharing Schemes, shared secret information cannot be revealed without any cryptographic computations. Various Secret Sharing Schemes have been proposed. However, the size of the shares and implementation complexity in these schemes depend on the number of participants. In other words, when a great number of participants are involved, the scheme will become impractical. A secret sharing scheme is called efficient if the total length of the n shares is polynomial in n: In the traditional Visual Secret Sharing Schemes, shared secret information can be revealed without any cryptographic computations.

[3] New Visual Secret Sharing Schemes Using Probabilistic Method

In this paper C. N. Yang proposed new methods to construct the VSS schemes. He uses a different approach, the probabilistic method. The major difference between these two schemes is that this scheme uses pixel operation and the conventional scheme uses sub pixel operation. Their ''OR''-ed operation of pixel is the same as the stacking operation of sub pixel m in the conventional VSS scheme. Instead of expanding the pixel into m sub pixels, they only use one pixel to represent one pixel. They use the frequency of white pixels to let human visual system distinguish between ''black'' and ''white''.

The new schemes have non-expansible shadow size and the same contrast level of the conventional VSS scheme. Here they use the term ''probabilistic'' to point out that our visual system distinguishes the contrast of the recovered image based on the difference of frequency of white color in black and white areas.

Algorithm 1: SA-based algorithm for GAS solver:

Procedure GAS_solver (n,n'max, ┌qual, ┌forb)
1. .n'←n
2. **Call GAS_SA**(n,n'max, ┌qual, ┌forb,Cbest,Zbest)
3. **If Zbest>+1 then** //No solution in last turn
4. n'←n'+1
5. **IF n'=n'max then Stop** and **Output** "No solution found"
6. **If n'<n then goto** Step 11
Else
7. .C←Cbest //Found a solution in last turn
8. **If n'>n then goto** Step 11
9. .n'←n'-1 //Improvement
10. **End If**
11. **Goto** Step 2
12. **Output n',C**, .

Algorithm 2: Pseudocode of Proposed SA-based Algorithm

Procedure **GAS_SA**(n,n'max, ┌qual, ┌forb,Cbest,Zbest)
1. ¥, *1<=i<=n,1<=j<=n',xij=1*
2. Calculate Ep for the above initial guess.
3. Zold=Ep,zbest=Zold,Cbest=C.
4.*t=t0,r=r0* .
5. **While Zold>=1 and t>=tf do**
6. **Repeat r times**
7. Randomly select a share ci .
8. Alter the solution configuration of ci randomly.
9. Calculate Ep for the new configuration.
10. .Znew=Ep
11.Δ E=Znew-Zold .
12. Generate a random number uniformly distributed in [0, 1).
13. **If ΔE<0**or p<e^(-Δ E/t) **then**
14. .Zold=Znew
15. **If Zold<Zbest then Zbest=Zold, Cbest=C** ,
16. **If Zold<1 then goto** Step 21

17. **else** recover the action in Step 8.
18. **End Repeat**
19 $t=\alpha\ t*t, r=\beta\ t*r$
20. **End While**
21. **Output** .Cbest,Zbest

Qual=qualified set, forb=forbidden set, Zbest =best found energy function,
Cbest =construction set best, Ep=penalized energy function

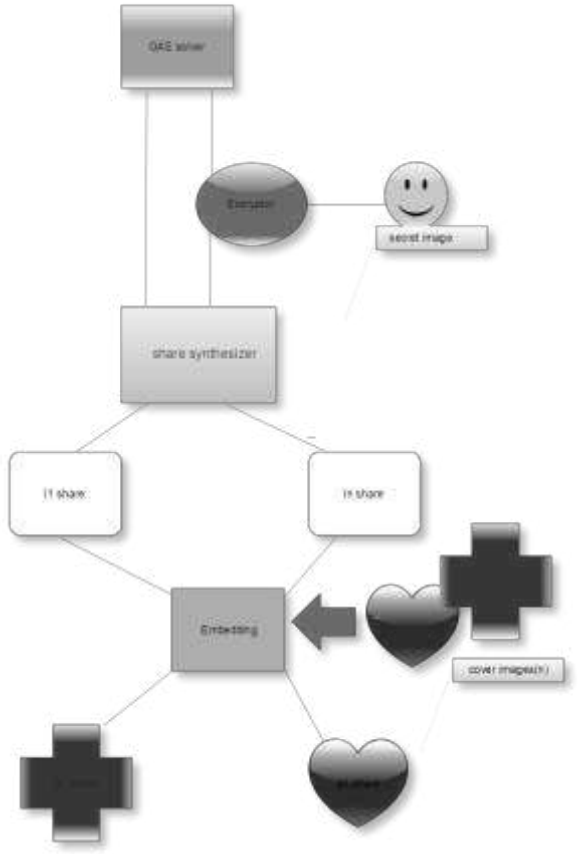
Algorithm 3: The Embedding Process:

Input: The n covering shares the corresponding VCS (C0, C1) with pixel expansion m and the secret image I.
Output: The n embedded shares e0,e1….en-1.

Step 1: Dividing the covering shares into blocks that Contain t(>=m) subpixels each.
Step 2: Choose *m* embedding positions in each block in the n covering shares.
Step 3: For each black (respectively, white) pixel in I, randomly choose a share matrix M
Step 4: Embed the subpixels of each row of the share matrix M into the *m* e mbedding positions chosen in Step 2.

ARCHITECTURE DIAGRAM

Figure-2



Sources: Authors Compilation.

CONCLUSIONS

Visual Cryptography is an image encryption technique used to hide the secure information in images. It allows the encryption of secret image into n number of shares and distributed into n number of participants. For example in (k, n) secret sharing problem the secret image can be visually recover by stacking together any k or more transparencies of the shares, but cannot reveal any secrete information by stacking less than k transparencies together.

The Embedded EVCS is constructed by adding random shares of secret image into meaningful covering images. The proposed system is worked with embedded EVCS approach for GAS which produces high quality sharing images. It also improves the contrast of the recovered secret image and produce clear resultant image.

REFERENCES

[1]. Embedded Extended Visual Cryptography Schemes F Liu… - … Forensics and Security, IEEE Transactions on, 2011 - ieeexplore.ieee.org

[2]. Z., Zhou; G., R. Arce, and G., Di Crescenzo, (2006). "Halftone Visual Cryptography", IEEE Trans. Image Process, August, Volume: 15, Number: 8, pp. 2441–2453.

[3]. M., Nakajima, and Y., Yamaguchi, (2002). "Extended Visual Cryptography for Natural Images", Proceedings WSCG Conference 2002, pp. 303–412.

[4]. D., S. Tsai; T., Chenc, and G., Horng, (2008). "On Generating Meaningful Shares in Visual Secret Sharing Scheme", Imag. Sci. J., Volume: 56, pp. 49–55.

[5]. Z., M. Wang; G., R. Arce, and G., Di Crescenzo, (2009). "Halftone Visual Cryptography via Error Diffusion", IEEE Trans. Inf. Forensics Security, September, Volume: 4, Number: 3, pp. 383–396.

[6]. M., Naor, and B., Pinkas, (1997). "Visual Authentication and Identification", Proceedings CRYPTO'97, Volume: 1294, pp. 322– 336, Springer-Verlag LNCS.

[7]. T., H. Chen, and D., S. Tsai, (2006). "Owner-Customer Right Protection Mechanism Using a Watermarking Scheme and a Watermarking Protocol", Pattern Recognit., Volume: 39, pp. 1530–1541.

[8]. P., Tuyls; T., Kevenaar; G., J. Schrijen; T., Staring, and M., Van Dijk, (2004). "Security Displays Enabling Secure Communications", Proceedings First International Conference Pervasive Computing, Boppard Germany, Springer-Verlag Berlin LNCS, Volume: 2802, pp. 271–284.

[9]. M., Naor, and A., Shamir, (1995). "Visual Cryptography", Proc. OCRYPT' 94, Berlin, Germany, Volume: 950, pp. 1–12, Springer-Verlag, LNCS.

[10].L., A. MacPherson, (2002). "Grey Level Visual Cryptography for General Access Structures", Master Thesis, University of Waterloo, Waterloo, ON, Canada.

****

INVESTIGATION OF BEARING FAULTS USING LAB VIEW

M. Senthil Kumar[107] S. Sendhil Kumar[108]

ABSTRACT

Vibration is a dynamic phenomenon which varies with time. The problems encountered in rotating machineries are of great concern to designers and maintenance engineers. Most failures of rotating machinery are rooting in the damage of rolling element bearings. The widespread applications of rolling element bearings in both industry and commercial life require advanced technologies to efficiently and effectively monitor their health status. In practice, it is very difficult to avoid vibration during machine running conditions. To know the level of damage on bearing during machine running conditions, vibration measurement should be carried out by suitable sensors / transducers. Accelerometer transducers are used almost in all vibration measurements. Many problems arising in motor operations are linked to bearing faults. In this paper, bearing vibration frequency features are discussed for motor bearing fault diagnosis. This paper presents an approach for rolling element bearing fault diagnosis using neural networks and time / frequency - domain vibration analysis. Vibration simulation is used to assist in the design of various motor rolling bearing fault diagnosis strategies. Both simulation and real-world testing results obtained indicate that can be effective agents in the diagnosis of various motor bearing faults through the measurement and interpretation of motor bearing vibration signatures.

KEYWORDS

Bearing, Vibration, Fault Diagnosis, Frequency, Simulation, Sensor, Motor, Domain, Accelerometer Transducers etc

1. INTRODUCTION

The problems encountered in rotating machineries are of great concern to designers and maintenance engineers. It has been observed by the various researchers on several occasions that rotating machinery stability conditions can change the operating conditions between the driver and the driven machines. Due to the high speed of some rotating machinery, the need for a better understanding of the phenomena is becoming a necessity for practical engineers for the purpose of troubleshooting. Most rotating equipment consists of a driver and driven machine coupled through a mechanical coupling. The mechanical coupling is used mainly to transmit torque from the driver to the driven machine. The two connected machines can go under misalignment, shaft bow, unbalance, etc., where severe misalignment can cause high vibration to the machinery assembly. Also, condition monitoring (CM) equipment is used extensively in the energy, petrochemical, cement, steel, paper, and pulp industries.

In addition, these industries are constantly under economic pressure to reduce costs while increasing service and productivity. Hence, CM is increasingly becoming an important asset management tool. Consequently, end users are looking to online, continuous CM systems to replace fixed-interval scheduled maintenance practices.

In today's industry, maintenance is one of the main procedures which has considerably high amount of cost regularly. Due to the amount and regularity of the cost, maintenance has been paid a long and dedicated attention by researchers and engineers since the start of automated mass production in industry. The most of the maintenance budget was spent on ineffective methods.

There are apparent reasons of ineffectiveness, which can be given alongside the definitions of methods. Those methods, which have been used in industry so far in order to reduce the cost, can be grouped into categories, which are worth mentioning. Different definitions and grouping can be found in different resources, but maintenance methods are here divided into three: a) run-to-failure, b) preventive, and c) predictive maintenance.

Predictive maintenance technique helps to determine the condition of in-service equipment in order to predict when maintenance should be performed. This approach offers cost savings over routine or time-based preventive maintenance because tasks are performed only when warranted.

II. BEARING CONDITION MONITORING

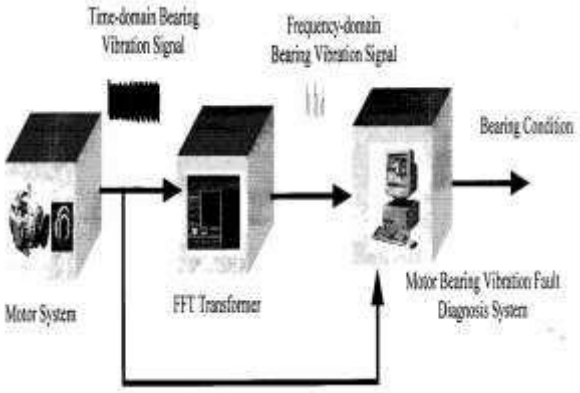
Due to the close relationship between motor system development and bearing assembly performance, it is difficult to imagine the progress of modern rotating machinery. In addition, the faults arising in motors are often linked with bearing faults. In many instances, the accuracy of the instruments and devices used to monitor and control the motor system is highly dependent on the dynamic performance of bearings. Bearing vibration can generate noise and degrade the quality of a product line which is driven by a motor system.

Heavy bearing vibration can even cause the entire motor system to function incorrectly, resulting in downtime for the system and economic loss to the customer. Proper monitoring of bearing vibration levels in a motor system is highly cost effective in minimizing maintenance downtime both by providing advance warning and lead time to prepare appropriate corrective actions, and by ensuring that the system does not deteriorate to a condition where emergency action is required. Thus, it is important to include bearing vibration diagnosis into the scheme of motor system fault diagnosis.

Signals from vibration sensors are usually measured and compared with reference measurements in order to interpret bearing conditions. The methods used to analyze these signals include probabilistic analysis, frequency analysis, time-domain analysis, and finite-element analysis. Among these methods, the frequency analysis approach is the most popular one. This popularity is most probably due to the availability of Fourier transform technique, as characteristics of vibration signals are more easily noticed in the frequency domain rather than in the time domain. The frequency analysis technique involves frequency analysis of vibration signal and further processing of the resulting spectrum to obtain clearly defined diagnosis information.

[107]Research Scholar, PSG College of Technology, Tamil Nadu, India, ssk724@gmail.com
[108]Research Scholar, Karpagam Institute of Technology, Tamil Nadu, India, ssk333c@yahoo.co.in

Figure-1: A Typical Motor Bearing Fault Detection Process



Sources: Authors Compilation.

Figure-2: Experimental Setup

Sources: Authors Compilation.

III. DATA ACQUISITION

With data acquisition, we take the first steps into the domain of practical vibration analysis. It includes following main tasks:

- Collection of machinery vibration,
- Conversion of the vibration signal to an electrical signal,
- Transformation of the electrical signal to its components,
- Providing information and documentation related to vibration data.

The above entails the entire hardware of the vibration analysis system Fig1 & 2. It includes transducers, electronic instruments that store and analyze data, the software that assist in vibration analysis, record keeping and documentation. To measure machinery or structural vibration, a transducer or a vibration pickup is used. A transducer is a device that converts one type of energy, such as vibration, into a different type of energy, usually an electric current or voltage. Commonly used transducers are velocity pickups, accelerometers and eddy current or proximity probes. Each type of transducer has distinct advantages for certain applications.

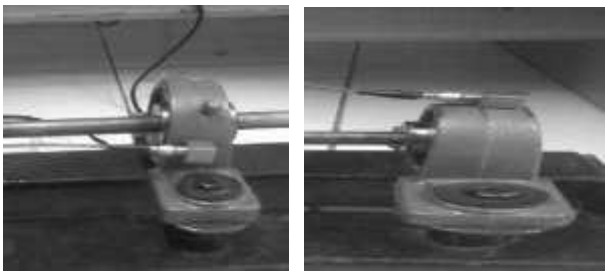
IV. ACCELEROMETER

Accelerometers are inertial measurement devices that convert mechanical motion into a voltage signal. The signal is proportional to the vibration's acceleration using the piezoelectric principle. They are rugged, compact, lightweight transducers with a wide frequency response range. Accelerometers are extensively used in many condition-monitoring applications. Components such as rolling element bearings or gear sets generate high vibration frequencies when defective. Machines with these components should be monitored with accelerometers as shown in Fig 3&4. When the accelerometer is subjected to vibration, mass exerts a varying force on the piezoelectric crystal, which is directly proportional to the vibratory acceleration. The charge produced by piezoelectric crystal is proportional to varying vibratory force.

V. BEARING VIBRATION SIGNATURE

As per design a model is fabricated with proper conditions, which is used to measure the vibration signals. In test, ring accelerometer is placed in the bearing mounting and frequency response, velocity rms value, displacement value has been taken. The vibration signals have been taken from coupling end as shown in fig 3.

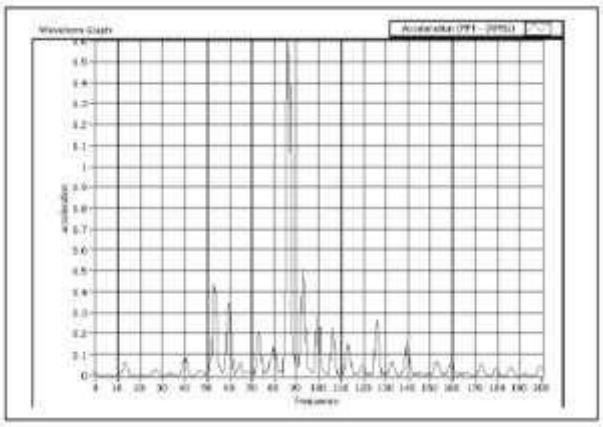
Figure-3: Positions of Accelerometer



Sources: Authors Compilation.

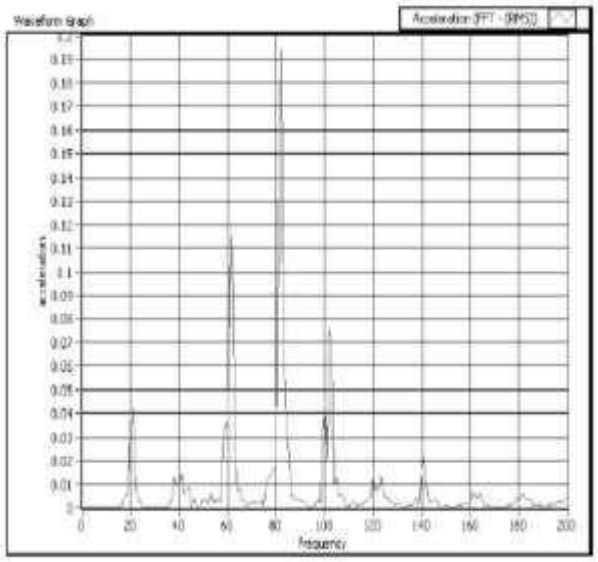
Frequency spectrum has been taken at the coupling and non-drive end bearing for three different speeds (750 RPM, 1000 RPM and 1250 RPM). The frequency spectrum for 0 mm coupling, mass and unbalance placed at middle from coupling end and non-drive end. Frequency spectrum of coupling end bearing is shown in Fig 4, 5, 6. Frequency spectrum measured from Non-drive end bearing (0 mm coupling) at various speeds (750, 1000, 1250 RPM) is shown in Fig 7, 8, 9.

Figure-4: Acceleration spectrum at a speed of 750 rpm



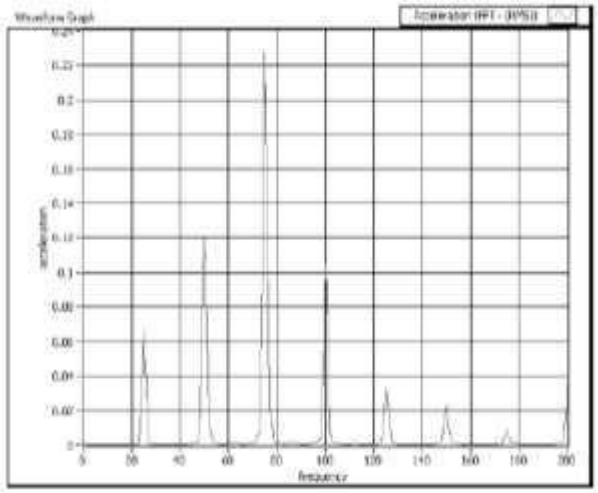
Sources: Authors Compilation.

Figure-5: Acceleration Spectrum at Speed of 750 rpm



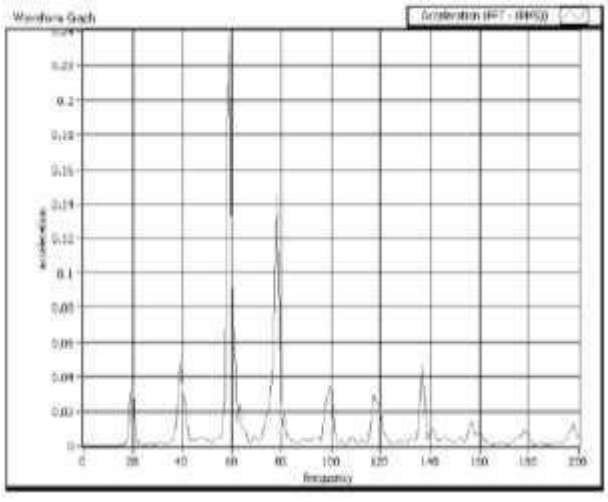
Sources: Authors Compilation.

Figure-6: Acceleration Spectrum at Speed of 1250 rpm



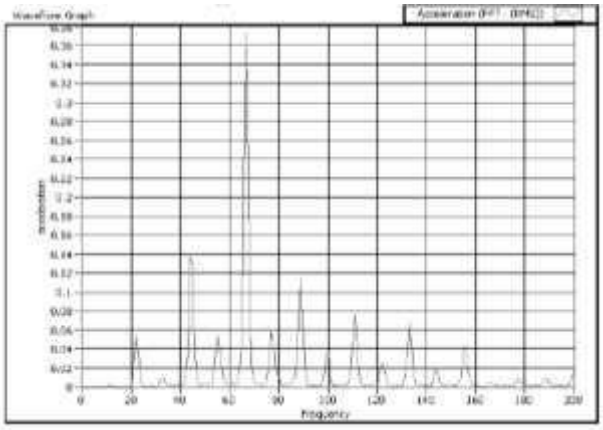
Sources: Authors Compilation.

Figure-7: 750 RPM



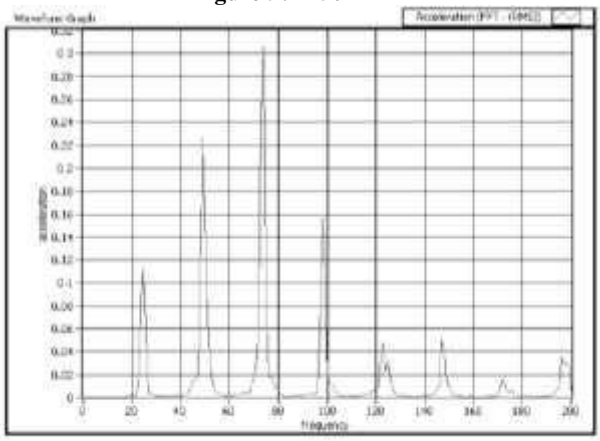
Sources: Authors Compilation.

Figure-8: 1000 RPM



Sources: Authors Compilation.

Figure-9: 1250 RPM



Sources: Authors Compilation.

V. RESULTS AND DISCUSSIONS

5.1 Displacement Spectrum

The displacement plot for various unbalance & configuration of location of the mass the spectrum is plotted & tabulated. The frequency spectrum has been taken from coupling end and non drive end also. Maximum value of displacement has been tabulated in the Table1.

From the values as the unbalance varies the value of displacement predominantly increases.

5.2 Velocity RMS Value

The RMS velocity is measured for the various configuration and location of the mass at middle. The RMS (Velocity) has been taken from coupling end and non drive end. Maximum value of velocity has been tabulated in Table2.

5.3 Acceleration Spectrum

The various acceleration spectrums are illustrated in Fig4–9 for various speeds. Fig 4-6 illustrates the spectrum at the coupling end bearing location. Fig 7-9 illustrates the spectrum at the non drive end location.

VI. CONCLUSIONS

By using the vibration severity charts refer the coupling end and non-drive end bearing vibration velocity rms value indicate the good, satisfactory, unsatisfactory and unacceptable region clearly.

VII. FUTURE WORK

Evaluate the health condition of the bearing, Develop a neural-network-based model for predicting bearing failure. Model the vibration signals using artificial neural networks to estimate failure times of a degrading bearing in order to make better decisions regarding scheduling of maintenance activities and replacement strategies.

VIII. REFERENCES

[1] Sebastian, Villwock, and Henning, Zoubek, (2007). "Rolling Bearing Condition Monitoring Based on Frequency Response Analysis", Institute of Power Electronics and Electrical Drives Holderlinstrasse 3, IEEE 2007, Volume: l, Number: 4.1.0

[2] Wei, Zhou; Thomas, G. Habetler, and Ronald, G. Harley, "Bearing Condition Monitoring Methods for Electric Machines: A General Review", 1-4244-1062-2/07/ ©2007 IEEE.

[3] Butterworth–Heinemann, (1999). "Root Cause Failure Analysis & Vibration Fundamentals", R. Keith Mobley Boston Oxford: A Member of Reed Elsevier Group 1999 by Butterworth–Heinemann, Vibration fundamentals book part2.

[4] T., Hoshi, (2005). "Damage Monitoring of Ball Bearing", Technical Research, Toyohashi, Japan.

[5] Runqing, Huanga; Lifeng, Xia; Xinglin, Lib; C., Richard Liuc; Hai, Qiud, and Jay, Lee, (2005). "Residual Life Predictions for Ball Bearings Based on Self-Organizing Map and Back Propagation Neural Network Methods", a Department of Industrial Engineering & Management, Shanghai Jiao Tong University ''Received 23 January 2005; accepted by 21 November 2005.

[6] Changting, Wang, and Robert, X. Gao, (2000). "A Virtual Instrumentation System for Integrated Bearing Condition Monitoring", Transactions on Instrumentation and Measurement, Volume: 49, Number: 2, April.

*****

CLUSTERBASED ENERGY EFFICIENT AND APPROXIMATE DATA COLLECTION FOR WIRELESS SENSOR NETWORKS

Dhivya S.[109] Vinoth Selvin S.[110]

ABSTRACT

In Wireless Sensor Networks, unambiguously sending huge amount of data to the destinationleads to several problems. So, the data aggregation is an essential technique in wireless sensor network. Certain application needs to collect data approximately.In this paper, we focus on to collect a data approximately and efficiently with limited energy and also needs to collect data securely. So, we insinuate a technique called ADC (Approximate Data Collection), which divides sensor network into cluster and determine data organization on each cluster head and the data will be approximated at the sink node. In specific, we offer an inception algorithm to approximate the sensor node readings in cluster head. Now-a-days sensor networks are used for security related applications and security is also become main anxiety. The proposed technique is adaptive to environment changes, more efficient and the amounts of message transmission also get reduced when compared to existing techniques.

KEYWORDS

Wireless, Sensor, Networks, Data, Aggregation, Inception, Algorithm etc

INTRODUCTION

The wireless sensor networks consist of huge number of sensor nodes which are densely deployed. Sensor network are deployed in an open space. The sensor networks are used in applications such as military surveillance, battle fields in front of or behind enemy lines, chemical industry etc. A wireless sensor networks consist of limited resource bounded sensor, which is equipped with inadequate battery power, so the transmission may consumes a lot of energy. To overcome this problem, data aggregation technique is taken into an account. This technique efficiently eliminates data redundancy and consumes energy. In many applications of wireless sensor network, it is not necessary to collect complete data set from a resource bounded WSN. In wireless sensor networks, directly sending large amount of data to sink node may cause several undesirable problems. The problems occur here are the quality of data gets lost because of packet loss, rigorous data collection leads to excessive communication overhead. Due to this the sensor nodes lifetime gets deployed. So the data collection strategy must be designed carefully to minimize energy consumption and to increase the lifetime of the network as much as possible.

For long term data collection in wireless sensor networks with limited bandwidth, approximate data collection (ADC) is the best method. A recent application of WSNs needs to collect data approximately and efficiently due to constraints in energy budget and communication bandwidth.

[109] PG Scholar, Karunya University, Tamil Nadu, India, dhiv.dhivs@gmail.com
[110]Assistant Professor, Sri Dhanalakshmisrinivasan College of Engineering, Tamil Nadu, India, vinothselvin@gmail.com

In this paper we consider the problem of data collection in physical environments, where the eachsensor containing a data are correlated. The correlation of the data being collected can be influenced by approximately fusing the data inside the network and reduces the number of transmission and energy consumption, for data collection process. To achieve this we propose a technique called ADC (Approximate Data Collection). Approximate Data Collection is a good choice for long term wireless sensor networks. In many applications with densely deployed sensor nodes, the sensor data gathered by sensor nodes may have a spatio-temporal correlation characteristic. By surveying such characteristics, sensor nodes collect their data in an aggressive manner. So the data traffic will be reduced by increasing the data accuracy. The Approximate Data Collection scheme is scalable and more efficient to environmental changes. The aggregated results are more important to make critical decisions and the accuracy of final aggregated result is more significant.

RELATED WORKS

There have been many related works on data collection in WSNs.The Ken method in wireless sensor network is a general data collection mechanism. It is a robust approximate technique. It uses a dynamic probabilistic model to reduce the communication between sensor node and base station. This method is also well suited to detect anomaly, in addition to data collection. The dis-advantage of using this technique is robust to communication failure that is, if the data lost during transmission it cannot be recovered. The advantage of this technique is that the readings which are visible to the user are with in fixed error bound [2]

The directed diffusion is a data centric approach, here all the communication is for the data which are named. In directed diffusion method sink broadcast its request to its neighbor. The neighbor cached the interests and the neighbor data which matching to the interest will respond back to sink. In this scenario the transmission will be reduced because of interested neighbor will be responded. The disadvantage of using this method, it's not suitable for long term large scale wireless sensor networks [3]. Cougar andTiny DB [12], [13] provide query-based interfaces to extractdata from sensor networks.

Those works mainly focus onquery-based data gathering, but none of them considerthe case of efficient long-term large-scale data collection.The similarity based adaptive framework method reduces the amount of communication in wireless sensor networks. A SAF method detects data similarities among the nodes and groups them into clusters. Its main goal is to conserve energy by reducing transmission. The disadvantage of using this method is, it's only taking an advantage of temporal correlation with in the sensor data without considering the similar readings of nearby sensor nodes [4].

The data collection approach called BBQ (Bar – Q – Method). It is a query based approach, which is used to extract data from sensor networks. It is the first method which uses Gaussian joint distribution model to extract data from sensor network. It will find correlations among sensor readings by using the data extracted from the sensor networks. It has several drawbacks because of using Gaussian joint distribution model. First it is not

suitable for large scale wireless sensor networks. Second, it needs to gather complete data set, gathering of complete data set is too energy consuming [8].

Another existing approach called Constraint Chaining. CONCH is a continuous monitoring scheme. This scheme uses a spatial and temporal correlation to attain energy saving. In a sensor network, the monitored values are either spatially or temporarily correlated, it suggest that it's not needed to report every value in every time if it is not changed from the existing value. It will assign reporter and updater at the edge, the role of the reporter is to report the base station if there is a change in value and the updater job is to update the value reported by the reporter. The disadvantage of using this technique is that it goes through update message overhead when the readings changes often [11].

DESIGN OBJECTIVE

Our scheme's design objective is to accomplish accuracy in data aggregation with limited communication overhead. To achieve a desired data aggregation scheme, the following criterion has to be satisfied:

Efficiency: To achieve efficiency, overhead has to be reduced.

Accuracy: The accuracy should be taken as an important criterion to evaluate the performance and it is used to make critical decisions.

Privacy: To preserve privacy against eaves dropping.

PROPOSED TECHNIQUE

In wireless sensor network, communication overhead and energy consumption due to packet losses is the major problem. To avoid those, various data aggregation schemes has been used. In this paper we focus on to reduce the packet loss and energy consumption, we proposes technique called approximate data collection and to approximate the reading sensed by a sensor node we propose an algorithm called inception algorithm.

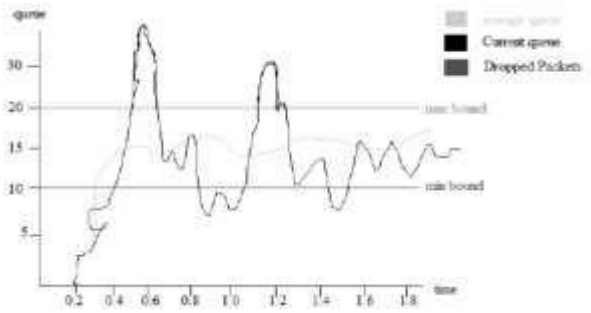
Inception Algorithm

The inception algorithm is similar to that of RED (Random Early Detection) Gateway and Doorway Algorithm [6]. This algorithm is mainly usedto reduce the congestion and also the transmission cost among sensor nodes and cluster head. The inception algorithm will control the queue size and marking packets by avoiding the congestion.

To generate a new sensor reading from its historic data the local estimation model [5] is utilized. The sensor node sends each nodes parameter to cluster head, instead of sending a raw data. The inception algorithm is mainly used to calculate the average size of the queue. The average size of the queue is compared with two threshold value it sets, that is maximum bound value and minimum bound value. Based on that criteria it will decide whether to transmit or to drop. The following scenarios are:

- If average queue size is less than the minimum bound value, it will not mark any of the packets it receives.
- If average queue size is greater than the maximum bound value, it will mark every packet and eliminates it.
- If average queue size is between the minimum bound value and maximum bound value, the arrived packets are marked with the probability '*pb'*.

Graph-1: Graph for Finding Queue Limit



Sources: Authors Compilation.

The inception algorithm is given by,

1. Initialize avg = 0, count = 1;
2. **For each** packet receives calculate the average size of the queue 'avg'
3. If the queue is not an empty queue **then** calculate
4. avg = (1 – wtq) * avg + wtq * cqs
5. **Else** n = l (t_cur – t_start); avg = (1 – wtq) ^ n * avg;
6. **end**
7. **if**mnbnd<= avg<mxbnd
8. count ++ **then** calculate
9. probability of 'pq'
   pb = mxpb* (avg – mnbnd) / (mxbnd – mnbnd);
   pq = pb / (1 – count * pb)
10. **end**
11. **else if**mxbnd<= avg then
12. mark the receiving packets
13. count = 0
14. **end**
15. **else** count = -1
16. **end**
17. **end**
18. **when** the queue becomes empty t_start = t_cur
19. **end**

Table-1: Notations

Notations	Description
avg	Average queue size
wtq	Queue weight
cqs	Size of the current queue
t_cur	Current time of the queue
t_start	Time the queue starts
mnbnd	Minimum bound value
mxbnd	Maximum bound value

Sources: Authors Compilation.

The advantage of using inception algorithm is that, the user visible readings will be guaranteed to be with fixed error bound on the detection of anomaly. Because it will transmit the readings that are within a threshold limit that is minimum bound and maximum bound value.

Approximate Data Collection (ADC)

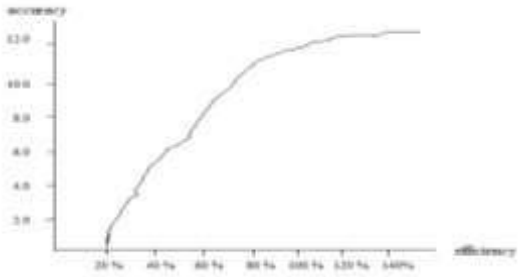
Wireless Sensor Network consists of huge number of sensor nodes. The sensor nodes will be deployed in an area to detect and to monitor some events. When the sensor detects an event that is to be monitored, the sensor node will report the event to its base station. Then by receiving information about that event the base station will take an appropriate action. The major

problem in wireless sensor networks is communication overhead and energy consumption. The data collection is a fundamental task to overcome this problem. In many scenarios, the collection of complete data is not necessary from resource bounded wireless sensor network. In wireless sensor network, sending huge amount of data to sink may cause several problems such as packet loss, energy consumptionand network congestion. To avoid these problems, the technique called Approximate Data Collection is proposed. For endure data collection approximate data collection is a wise choice due to its limitations in energy budget and communication bandwidth. While designing an ADC (Approximate Data Collection) there are some factors to be considered. They are it should be scalable, efficient and it should be capable of adjusting its parameters automatically according to its environmental changes (self – adaptive).

By using ADC, the sensor network is divided into and each cluster have its own cluster head and it find data coordination on cluster heads, then by using an parameter updated by an data model it performs approximation on sink node. The cluster head will be chosen among the cluster members. The node which has more energy will be chosen as a cluster head. The cluster head will be selected by cluster members. The inception algorithm will be deployed into the cluster headto approximate the readings of sensor nodes. Each sensor node transmits its parameter to its cluster head, instead of intense sensor readings. The approximated value is compared with the original value, if the difference between them is not greater than the sensor node does not update its data to cluster head. If the difference between the original value and approximated value exceeds the threshold limit, the sensor node will sent an update message. As a result, the communication cost will be reduced.

SIMULATION RESULT

Graph-2: Accuracy and Efficiency for Inception Algorithm



Sources: Authors Compilation.

Graph2 shows the graph for a proposed inception algorithm. When compared to the existing schemes, the inception achieves higher accuracy and efficiency.

CONCLUSIONS

In wireless sensor networks, sensor nodes are battery powered. Communication is more power consuming than computation. Communication overhead is a major concern in wireless sensor networks. Energy consumption and communication overhead can be reduced and the lifetime of wireless sensor networks can be extended. Data aggregation accuracy is more important for some applications in WSNs; we propose an Approximate Data collection to providean accurate data aggregation. A novel method generates an exact result and reduces amount of transmission in communication.

REFERENCES

[1] M., Li; Y., Liu, and L., Chen, (2008). "Non-Threshold Based Event Detection for 3D Environment Monitoring in Sensor Networks", IEEE Trans. Knowledge and Data Eng., December, Volume: 20, Number: 12, pp. 1699-1711.

[2] D., Chu; A., Deshpande; J., M. Hellerstein, and W., Hong, (2006). "Approximate Data Collection in Sensor Networks Using Probabilistic Models", Proceedings 22nd International Conference Data Eng. (ICDE '06).

[3] C., Intanagonwiwat; R., Govindan, and D., Estrin, (2000). "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks", Proceedings MobiCom.

[4] D., Tulone, and S., Madden, (2006). "An Energy Efficient Querying Framework in Sensor Networks for Detecting Node Similarities", Proceedings Ninth ACM International Symp. Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWiM '06).

[5] Chao, Wang; Huadong, Ma; Yuan, He, and Shuguagn, Xiong, (2012). "Adaptive Approximate Data Collection for Wireless Sensor Networks", IEEE Transactions on Parallel and Distributed Systems, June, Volume: 23, Number: 6.

[6] The RED Queing Discipline, http://opalsoft.net/qos/DS-26.htm

[7] R., Rajagopalan, and P., K. Varshney, (2006). "Data-Aggregation Techniques in Sensor Networks: a Survey", *IEEE Comm. Surveys & Tutorials*, Volume: 8, Number: 4, pp. 48–63.

[8] Deshpande, C.; Guestrin, S.; Madden, J.; Hellerstein, and W., Hong, (2004). "Model-Driven Data Acquisition in Sensor Networks", Proceedings 13th International Conference Very Large Data Bases (VLDB '04).

[9] You-Chiun, Wang, "Data Compression Techniques in Wireless Sensor Networks"

[10] Yujie, Zhu; Ramanuja, Vedantham; Seung-Jong, Park, and Raghupathy, Sivakumar, (2008). "A Scalable Correlation Aware Aggregation Strategy For Wireless Sensor Networks", Information Fusion, July, Volume: 9, Issue: 3.

[11] Silberstein, R. Braynard, and J., Yang, (2006). "Constraint Chaining: On Energy-Effcient Continuous Monitoring in Sensor Networks", Proceedings ACM SIGMOD International Confernce Management of Data (SIGMOD '06).

[12] S., Madden; W., Hong; J., M. Hellerstein, and M., Franklin, (2011). "TinyDB Web Page: http://telegraph.cs.berkeley.edu/tinydb.

[13] Y., Yao, and J., Gehrke, (2003). "Query Processing in Sensor Networks", Proceedings CIDR.

*****

EXTENSIBLE SECURITY SUITE FOR MOBILE AD HOC NETWORKS ROUTING

C. Narmatha[111] D. Malini[112]

ABSTRACT

MANETs (Mobile Ad-Hoc Networks) are gaining great importance in commercial, industrial, and military applications. However, many of the current MANET security systems do not cover all the security aspects. The methodology utilizes powerful authentication and user-adjustable encryptions based on digital certificate chaining and popular ciphers such as DES, AES, and RSA. This has a highly modular design, which allows additional security strength to be added on top of the existing framework. This attractive feature enables heterogeneous applications to utilize various levels of security routing on the same MANET. The performance is compared with conventional authentication and encryption for MANET routing using NS2 simulator.

KEYWORDS

MANET, Security Systems, Encryptions etc

INTRODUCTION

A mobile ad hoc network (MANET), sometimes called a mobile mesh network, is a self-configuring network of mobile devices connected by wireless links. Each device in a MANET is free to move independently in any direction, and will therefore change its links to other devices frequently. Each must forward traffic unrelated to its own use, and therefore be a router. The primary challenge in building a MANET is equipping each device to continuously maintain the information required to properly route traffic. Such networks may operate by themselves or may be connected to the larger Internet. MANETs are a kind of wireless ad hoc networks that usually has a routable networking environment on top of a Link Layer ad hoc network. They are also a type of mesh network, but many mesh networks are not mobile or not wireless.

The growth of laptops and 802.11/Wi-Fi wireless networking have made MANETs a popular research topic since the mid- to late 1990s. Many academic papers evaluate protocols and abilities assuming varying degrees of mobility within a bounded space, usually with all nodes within a few hops of each other and usually with nodes sending data at a constant rate. Different protocols are then evaluated based on the packet drop rate, the overhead introduced by the routing protocol, and other measures.

1.1 Routing Protocols in Mobile Ad Hoc Networks

Ad hoc On-Demand Distance Vector (AODV) Routing is a routing protocol for mobile ad hoc networks (MANETs) and other wireless ad-hoc networks. It is a reactive routing protocol, meaning that it establishes a route to a destination only on demand. In contrast, the most common routing protocols of the Internet are proactive, meaning they find routing paths independently of the usage of the paths. AODV is, as the name indicates, a distance-vector routing protocol.

1.2 Advantages

1. The main advantage of this protocol is that routes are established on demand and destination sequence numbers are used to find the latest route to the destination.
2. The connection setup delay is lower.

RELATED WORKS

There are two fields relevant to the composition of the security suite – encryption and authentication. Our work combines and modifies many of the existing works on these two fields. Thus, we will be presenting major related works on encryptions and authentications

PROPOSED AODV

Ad hoc On-Demand Distance Vector (AODV) Routing is a routing protocol for mobile ad hoc networks (MANETs) and other wireless ad-hoc networks. It is a reactive routing protocol, meaning that it establishes a route to a destination only on demand. AODV is, as the name indicates, a distance-vector routing protocol.

AODV avoids the counting-to-infinity problem of other distance-vector protocols by using sequence numbers on route updates, a technique pioneered by DSDV (AODV is capable of both unicast and multicast routing.

It includes a simple but yet secure authentication mechanism, and four user-adjustable encryption modes that each has different encryption strength. By choosing an appropriate mode, the suite allows MANET users to balance the level of security needed and the minimal performance required. This suite also lays the security foundation, which allows for future extension. We demonstrate our design concept on a popular MANET protocol – AODV (Ad-hoc On-demand Distance Vector), and use simulations to show the best case and worst case performance impact of the suite on the network.

3.1 Security Problems in MANETs

MANETs are much more vulnerable to attack than wired network. This is because of the following reasons. Open Medium - Eavesdropping is easier than in wired network. Dynamically Changing Network Topology - Mobile Nodes comes and goes from the network, thereby allowing any malicious node to join the network without being detected.

Cooperative Algorithms - The routing algorithm of MANETs requires mutual trust between nodes which violates the principles of Network Security.

Lack of Centralized Monitoring - Absence of any centralized infrastructure prohibits any monitoring agent in the system.

[111] Assistant Professor, Dhanlakshmi Srinivasan College of Engineering, Tamil Nadu, India, narmuchandran@gmail.com
[112]Lecturer, Dhanlakshmi Srinivasan College of Engineering, Tamil Nadu, Tamil Nadu, India, malini.gds@rediffmail.com

3.2 Challenges for Security

The various challenges for security in e-commerce have been listed below. They are popularly known as the five pillar s of e-commerce.

- Confidentiality,
- Authentication,
- Integrity,
- Non- repudiation,
- Interoperability / Universality.

3.2.1 Confidentiality

Confidentiality deserves the process of maintaining the information you are sending, such as credit card information when purchasing goods online, or sensitive business information in e-mail, that can't be read by anyone other than the intended recipient.

3.2.2 Authentication

Authentication deals with the process of verifying the identity of users. For example, if a person wouldn't want a competitor to download his company information from outside. As a user, user also may want to be sure that which part is authenticated for him for buying goods from an online store. It emulates the security which leads to keep the user away from credit card and internet hackers.

3.2.3 Integrity

Integrity makes the process more secured and unchanged. It swears that no one has intercepted information and changed it in any way. So tampering of the information by anybody should be difficult and evident.

3.2.4 Non-repudiation

In the real world, a contract with a written signature is generally binding. There is no real equivalent on the Internet. Someone might buy some stock over the Internet, the price falls, and then they say they never placed the order. There isn't a way to sign a contract electronically except with a certificate.

3.2.5 Interoperability

This suite also lays the security foundation, which allows for future extension. Our design concept on a popular MANET protocol – AODV (Ad-hoc On-demand Distance Vector), and use simulations to show the best case and worst case performance impact of the suite on the network.

There are two fields relevant to the composition of the security suite – encryption and authentication. Our work combines and modifies many of the existing works on these two fields. Thus, we will be presenting major related works on encryptions and authentications.

PROJECT GOAL

This project report presents an Advance security enhancement for MANETs, which obviously creates an advanced package for existing AODV and thus eliminates various security attacks. Our main goals in this project are:

Comparative study between Native AODV with the proposed system under three categories:

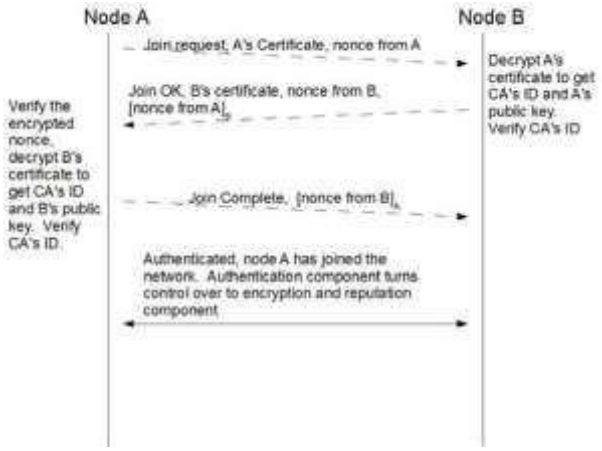
- Authentication,
- Encryption,
- Data Integrity.

METHODOLOGY

Authentication Scheme

Authentication protocols for wireless networks, but most of them are designed for infrastructure networks. The authentication scheme in digital certificate using distributed public cryptography has been proven to be reliable and has performed reasonably well. In authentication scheme and encryption schemes both use a modified version of digital certificate. Each node, before joining the mobile ad hoc network, is assumed to have a secure and reliable mean of communicating with a Central Authority.

Figure-1: Authentication Handshakes



Sources: Authors Compilation.

Node A is the joining node and node B is the existing node in the network. When node A first joins, it sends Join Request along with its certificate and a nonce. Node B, upon receiving the message, replies with a join ok that includes B's certificate, a new nonce, and nonce from A signed by B. Finally, A sends join complete with B's nonce signed by A. In this way, nodes contact the CA once in a while to obtain certificates. The neighboring nodes exchange certificates with a node upon its arrival to the network, and they verify the signature of the certificates against the well-known public key of the CA. Nodes establish a direct connection with another node only if the signature of the corresponding certificate is valid. To keep the freshness of the authentication, each certificate expires after a certain time, after which the node needs to contact the CA again to obtain a new certificate.

Encryption Scheme

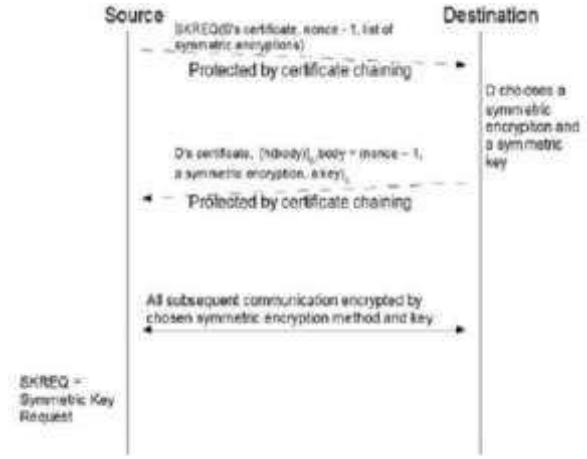
The encryption scheme in security suite relies on the certificate component of the authentication scheme. It is adjustable for achieving a balance between security and performance. It is based on distributed public cryptography. Security Suite provides four different modes of communication encryption for MANET. Each node using the security suite will have an Encryption Table of adjustable size. There are four modes of

encryption offered: Open Mode, User Mode, Lightweight Mode, and Strong Mode.

OPEN MODE: This mode has no encryption, and the data in the network are all readable. However, variety of attacks is still possible.

LIGHTWEIGHT MODE: In this mode, the encryption never expires, and so no further key exchanges are required after the first exchange.

Figure-2: Message Exchange to Establish Lightweight Mode Encryption



Sources: Authors Compilation.

STRONG MODE: This mode offers the strongest security in security suite. In this mode, the symmetric crypto used for data communication refreshes very quickly. Since the symmetric keys and methods change frequently, we can expect all the encryptions to expire and refresh before malicious adversaries have a chance to obtain enough meaningful data to begin cryptanalysis. To ensure the freshness and security strength, the key expires quickly. Each key in Strong Mode expires in the Minimum time of every 200 messages or one hour. After that, a new key and new symmetric encryption method will again be renegotiated.

Figure-3: Message Exchange to Establish Strong Mode Encryption

Sources: Authors Compilation.

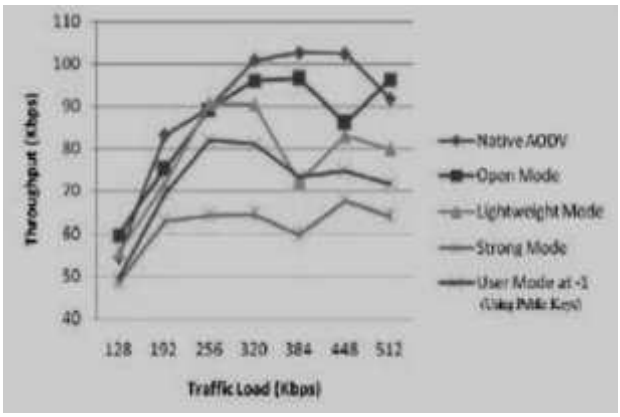
6. Performance Evaluation

Simulation is used to evaluate the impact of security suite on a Network performance. The Java-based simulator J-SIM is Chosen for the simulation .To evaluate the Performance, we simulate six AODV variants – The native AODV (with no authentication or encryption) and AODV with the four encryption modes: Open Mode, Lightweight Mode, Strong Mode Encryption, and User Mode.

In Light weight mode, a DES algorithm is used to generate a key of length 56 bits for encryption of data. The RSA key generation algorithm was used to generate a key of length 1024 bits for encryption of data in User Mode at refresh equals to -1 and Strong Mode. Strong mode, additionally, used AES algorithm to simulate the fast refreshing symmetric cipher.

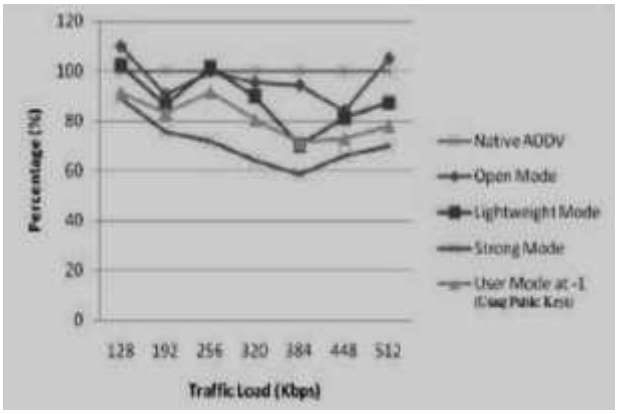
All the simulation runs have random mobility of 5~100meters per second in a pre-defined simulation area of 4 km by 4 km, and 50 nodes are present at any time during simulation. Each node has 802.11 Link-layer setting at 1 Mbps

Figure-4: Throughout.



Sources: Authors Compilation.

Figure-5: Throughput Ratio vs. Native AODV

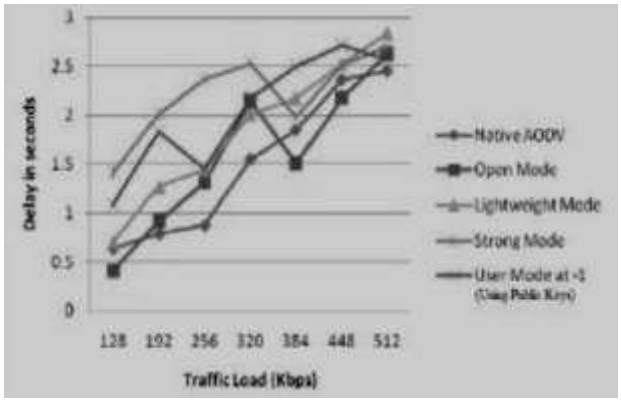


Sources: Authors Compilation.

Throughput is calculated as the average end-to-end Throughput of all transmission pairs. The throughput Performance results are shown in figure 4. Figure 5 shows the ratio comparison, normalized by the native AODV throughput. As shown in figure 5, Open mode performs very close to the native AODV with near 100% throughput. Lightweight mode follows closely with an achievable average of 90% native AODV throughput.

User Mode at -1 achieves around 80%, and Strong Mode achieves a tolerable but gapping 75% of native AODV throughput.

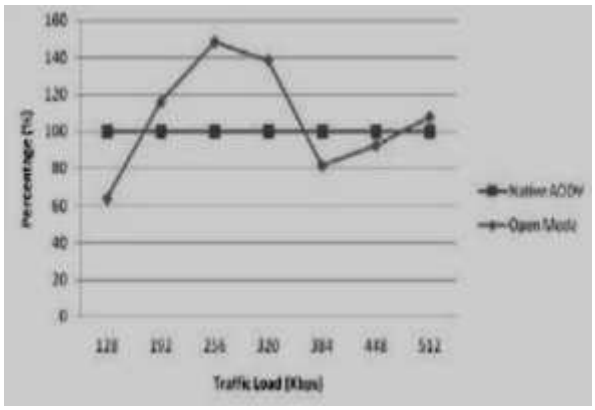
Figure-6: Delay



Sources: Authors Compilation.

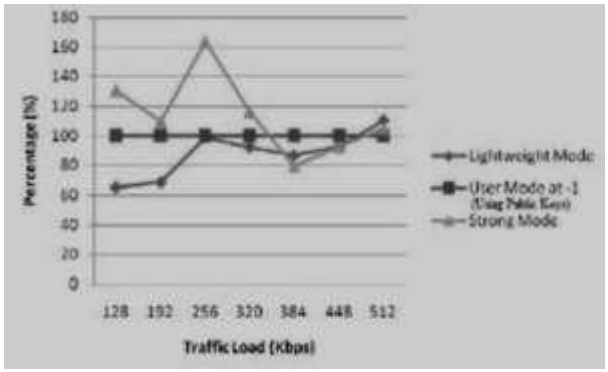
The delay performance is shown in Figure 6. Since J-SIM supports only a 1 Mbps link rate for IEEE 802.11 wireless links, the packet transmission time is high comparing with newer technologies, such as 56 Mbps (802.11g) or 150 Mbps (802.11n). This has caused high delay values, as shown in Figure 6, especially at high traffic loads.

Figure-7: Authentication Delay Ratio – Open Mode vs. Native AODV



Sources: Authors Compilation.

Figure-8: Encryption Delay Ratio – Lightweight and Strong Modes Vs. User Mode AT -1 (Using Public Key)



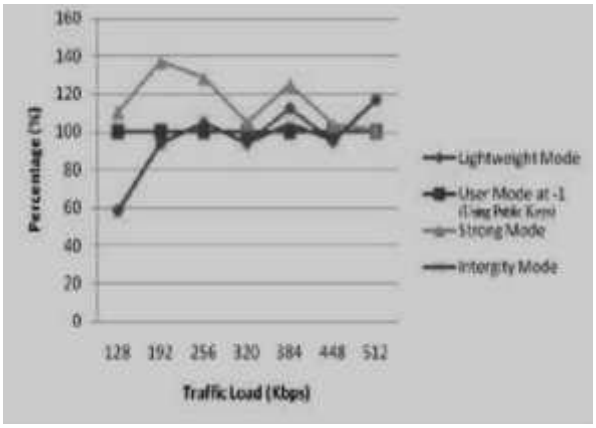
Sources: Authors Compilation.

In both throughput and delay performance, a slight dip in graph is observed at traffic load around 384Kbps. Comparison with other traffic models suggests this is attributed to the fixed randomness in J-SIM random mobility model.

Data integrity Mode Performace

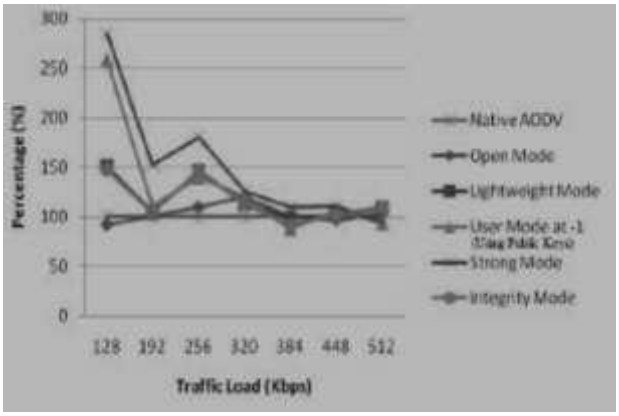
Data-integrity is simulated in the sixth variant of simulation. Figure 9 and figure 10 shows that, Data-integrity mode, by itself, performs closely to the performance of lightweight mode. Data-integrity mode performs worse than Open Mode due to the hashing and hash signing of the packets. Although not shown in figure here, when running in parallel with Lightweight Mode or Strong Mode, Data-integrity Mode only adds less than 1% more delay or throughput impact.

Figure-9: Encryption Delay Ratio Normalized to User Mode at -1 (Using Public Key) Under CBR Model



Sources: Authors Compilation.

Figure-10: Encryption Delay Ratio Normalized to Native AODV under CBR Model



Sources: Authors Compilation.

CONCLUSIONS

The presented Security Enhancement for MANETs is makes the configured nodes more secured. Hence, Even the strongest security can be achieved an average of 70-80% throughput while it moderately increasing the delay by 10-30%. Such a system can take advantage of certificate chaining to better monitor for malicious behaviors, and use certificate revocation feature in CA (Centralized authority) to keep compromised nodes out of the network. The main advantage of the Suite is in

using the user friendly Encryption as well as authentication rather than the conventional things.

REFERENCES

[1] Andel, Todd R, and Yasinsac, Alec., (2007). "Surveying SecurityAnalysisTechniques in MANETR Routing Protocols", IEEE Communications Survey, pp. 70-85.

[2] Carruthers, R., and Nikolaidis, I., (2005). "Certain Limitations of Reputation-Based Schemes in Mobile Environments", Proceedings of 8th ACM Int. Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems, pp. 2-11

[3] Dressler, F., (2005). "Reliable and Semi-reliable Communication with Authentication in Mobile adhoc Networks," Proceedings of IEEE International Conference on Mobile Adhc and Sensor Systems'05. pp. 5-6.

[4] Hu, R.Q.; Paranchych, D.; Mo-Han, Fong; Geng, Wu., (2007). "On the Evolution of Handoff Management and Network Architecture in WiMAX", IEEE Mobile WiMAX Symp., pp. 144-149.

[5] Hu, J., and Burmester, M., (2006). "LARS: a Locally Aware Reputation System for Mobile Ad Hoc Networks", 44th Annual Southeast Regional ACM Conference 2006. pp. 119-123.

[6] Kargl, F.; Schlott, S.; Kenk, A., and Geiss, A., "Securing ad hoc Throuting Protocols. 30", Euromicro Conference '04. pp. 514-519.

[7] Kassab, M.; Bonnin, J., and Belghith, A., (2008). "Fast and Secure Handover in Wlans: An Evaluation of the Signaling Overhead", IEEE Conference on Consumer Communications, January.

[8] Kassab, M.; Bonnin, J.; Belghith, A., and Sassi, S., (2005). "Fast Pre-Authentication Based on Proactive Key Distribution for 802.11 Infrastructure Networks", 1st ACM Workshop on Wireless Multimedia Networking and Perf. Modeling, pp 46-53.

[9] Kitada, Y.; Takemori, K.; Watanabe, A., and Sasase, I., (2005). "On Demand Distributed Public Key Management without Considering Routing Inf. and Telecommunication Technologies, pp. 375-380.

[10] Li, J.; Moh, T.-S., and Moh, M., (2009). "Path-Based Reputation System for MANET Routing", Proceedings 7th International Conference on Wired / Wireless Internet Communications (WWIC), Enschede, the Netherlands, May.

[11] Lopez, R. M.; Perez, G. M., and Gomez, Skarmeta A. F., (2005). "Implementing RADIUS and Diameter AAA Systems in IPv6- based scenarios", 19th International Conference on Advanced Information Networking and Applications, pp. 851-855.

[12] Mandin, Jeff, "802.16e Privacy Key Management (PKM) Version 2" IEEE 802.16 Broadband Wireless Access Working Group Project. http://ieee802.org/16

[13] Ngai, E. C. H., and Lyu, M. R., "An Authentication Service Based on Trust and Clustering in Wireless Ad Hoc Networks: Description and Security Evaluation", IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, pp. 94-103.

[14] Pi, Jian-Yong; Liu, Xin-song; Wu, Ai, and Liu, Dan, (2006). "A Novel Cryptography for Ad hoc Network Security", Proceedings of International Conference on Communications, Circuits and System '06, pp. 1448-1452.

*****

ENCRYPTION SCHEME BASED ON TURBO CODE USING DYNAMIC PUNCTURE MECHANISM

Mohammed Nawaz[113]

ABSTRACT

Turbo codes are the most powerful error correction code which enables effective and efficient error correction in different fields of communication. High secure data transmissions are carried out by encrypting the information in the original form. In most of the communication systems, error correction and information encryption are totally independent but encryption techniques are sensitive to noise. So in this paper proposes a new method of encryption based on Turbo code. To implement this method puncturing mechanism with a secret key is used in the turbo encoder which can be controlled by the normal random sequence. Puncturing mechanism deletes the parity bits randomly that enables the high error correction capability of the Turbo code. When decoding, only the legal receiver can generate the same normal random sequence using the secret key, then classify and decrypt the received sequence others get wrong decoding results if they tried to decode.

In this paper is implemented by using MatLab software here the security level of this method is checked by obtaining correlation result for various value of the code rate also obtain BER value for proposed encryption method to find the error-immunity. Finally this can be applied to the image encryption.

KEYWORDS

Turbo Codes, Puncturing, Correlation, BER, Image Encryption etc

INTRODUCTION

Turbo codes are one of the highly Advanced Error correction code which gives efficient and reliable output .When data is transmitted from transmitter to the receiver the main problem which affects is the level of the security. In order to provide Secrecy in communication and protect data from interception by illegal receivers in communication channels we use the technique Cryptography. There are different methods have been developed to encrypt and decrypt data in order to keep the information secret .The method such as as DES, RSA AES etc, have been formed to provide high security for information communication. The recent encryption techniques are very high sensitive to noise. So, the small quantity of noise in transmission causes the encryption system to collapse. To correct this problem we require an powerful code which can be used before transmission. We obtain both reliability and high security simultaneously by integrating error correction code with the encryption process in the communication system. The high security of the system is obtained when attackers will tried to decrypt the data with an wrong key they will obtain an incorrect different information. The reliability means the encryption system has a high level of immunity to channel errors; that is it must have a high level of error correcting capability.

[113] Student, Dhanlakshmi Srinivasan College of Engineering, Tamil Nadu, India, bfn786@gmail.com

In order to provide communication without errors we use the powerful error correction code Turbo code along with encryption process. Turbo codes emerged in 1993, Claude Berrou, Alain Glavieux and Punya Thitimajshima published the paper on which first described turbo codes. They showed that turbo codes could achieve low BERs at signal to noise ratios (SNR) very close to the Shannon Limit. It had previously been thought that the Shannon Limit on code performance could only be approached using codes with extremely long codeword lengths; i.e. codes which are potentially very complex to decode. Berrou et al. showed how the Shannon limit could be approached with realizable decoding complexity also made development by working on the Soft Output Viterbi Algorithm. Now it became a popular area of communications research.

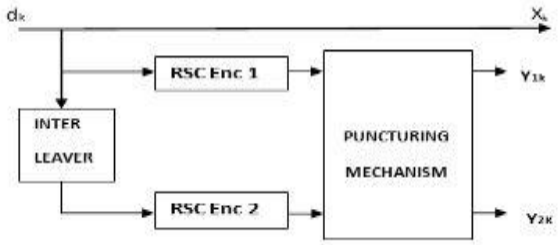
The excellent error correction performance which is obtained from Turbo code which is near to the Shannon limitation if the frame size is large enough. Therefore it is advantageous to design encryption scheme using Turbo code. The coding rate of the normal Turbo code is 1/3.The encoder for a turbo code consists of two convolutional codes in parallel, with their inputs separated by a pseudo-random interleaver. The decoder consists of two Maximum A Posteriori decoders connected in series via interleavers, with a feedback loop from the output of the second to the input of the first.

In this paper to obtain a higher coding rate, a puncturing mechanism is often adopted. By periodically eliminating some bits from the output of the recursive systematic convolutional encoders of the Turbo code, a higher coding rate can be achieved. Puncturing mechanisms delete the parity bits periodically. If we puncture the parity bits irregularly and control the puncturing algorithm with a secret key, the information will be encrypted during Turbo encoding. The main point of this dynamic puncturing mechanism is that the error correction capability of the Turbo code is ensured during the information encryption.

TURBO CODES AN OVERVIEW

The Turbo encoder which mainly consist of two blocks they are interleaver and two recursive systematic convolutional encoders RSC Enc 1 and RSC Enc 2. The RSC is the convolutional codes which use feedback and in which the uncoded data bits appear in the transmitted code bit sequence. In order to attain higher coding rate, the puncturing mechanism is adopted in the output of the two RSC encoders. Fig.1: shows the structure of the basic Turbo code.

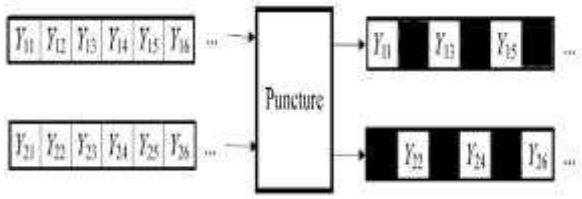
Figure-1: Structure of Normal Turbo Code



Sources: Authors Compilation.

In the puncturing mechanism of a normal Turbo code puncturing algorithm deletes the bits at even locations inY_{1k} and the bits at odd locations in Y_{2k}Where Y_{1k} in which the k is the k th output bit of RSC encoder1 ;where k = (1, 2, 3, 4,…) and Y2k in which the k is the k th output bit of RSC encoder 2; where k = (1, 2, 3, 4,…) , the deleted bits are usually located periodically. So the parity bits are reduced by half, and the coding rate of the Turbo code is increased to 0.5.

Figure-2: Demonstration of Puncturing in Normal Turbo Codes



Sources: Authors Compilation.

After completing the puncturing process, the transmitting sequence is $X_1, Y_{11}, X_2, Y_{22}, X_3, Y_{13}, X_4, Y_{24}, X_5, Y_{15}, X_6, Y_{26}$,…… where X_k in which k = (1, 2, 3, 4, …) is the k th bit of the first output of the encoder At receiver end the decoder uses same puncturing algorithm to classifyX_k, Y_{1k}, and Y_{2k} in the received sequence, and then sends them to a Turbo decoder to start an iterative decoding process. Here one parity bit is reserved for every k information bits.

The Bit Error Rate (BER) performances of the normal Turbo codes with coding rates are 1/2 and 1/3 respectively, when the SNR of the Additive White Gaussian Noise (AWGN) channel varies from 1.0 dB to 3.0 dB. In this experiment, the Turbo frame size is 400 bits, and the log-maximum a posteriori algorithm is implemented when decoding. In these simulation output w the coding rate increases from 1/3 to 1/2, the SNR increases about 0.8dB. This gives scenario for the turbo code.

PROPOSED TURBO BASED ENCRYPTION SCHEME

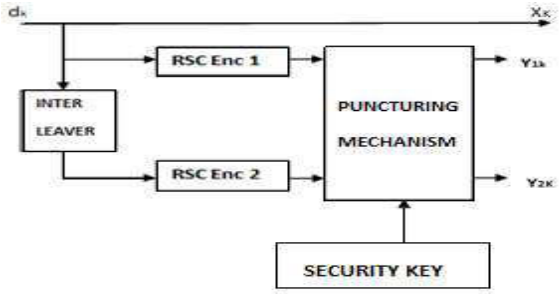
The proposed method of Encryption scheme is similar to that of the normal Turbo code. The consistency is necessary for the encoder and decoder of the Turbo coding scheme. Apart from normal turbo code here I adopt the dynamic puncturing Mechanism with a secret key to control it. The legal receiver who has the key can only able to classify X_k ,Y_{1k} and Y_{2k} correctly with the same puncturing mechanism, and then to decode successfully.

To ensure a low BER values, reserved parity bits should be irrelevant as much as possible. Any improper work in puncturing will automatically reduce the error correction capability. This proposed scheme enables good security and precise error correction capability with one coding step. Below shows detailed view of the encryption and decryption process basis of Turbo codes.

A. Encryption Process

The encoder part consist of the secret key that corrects the puncturing mechanism remaining blocks are almost same to that of the normal Turbo encoder. Fig.3: shows the structure of the turbo based encoder scheme section.

Figure-3: Block Diagram for Proposed Encryption Scheme



Sources: Authors Compilation.

Let the length of the RSC encoder is *K*, then corresponding memory will be *K*-1, and generators of the two RSC encoders are represented as:

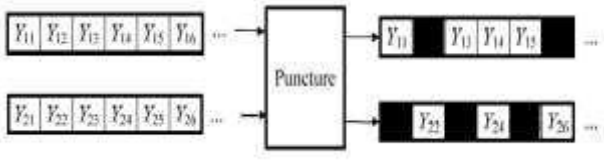
$G1 = [g_{10}, g_{11}, \ldots, g_1, K\text{-}1]$;
$G2 = [g_{20}, g_{21}, \ldots, g_2, K\text{-}1]$.

We can obtain various puncturing schemes by changing the secret key of the puncturing mechanism for our proposed encryption scheme. The changing of code rate from higher to lower which ensure a high error correction capability of coding scheme. To implement the turbo based encryption scheme which involves the certain algorithm to complete the process.

Below shows detailed view of this algorithm:

- Consider the secret key be s.
- By using *s* as an initial value, generate a normal random sequence *P* in which each elements are integers, the mean of the sequence is 0 and the standard deviation is *d*. The length of the sequence which is equal to the frame size of the Turbo code.
- If P(k) is 0 and i is even (k = 1, 2, 3, 4 …), delete the output bit of RSC encoder1 Y_{1k}; if P(k) is 0 and k is odd, delete the output bit of RSC encoder2 Y_{2k}.
- If P(k) is not 0, both Y_{1k} and $Y_{2k}$are reserved by this process, a key-controlled dynamic puncture is achieved.

Figure-4: Demonstration of Proposed Puncturing Scheme



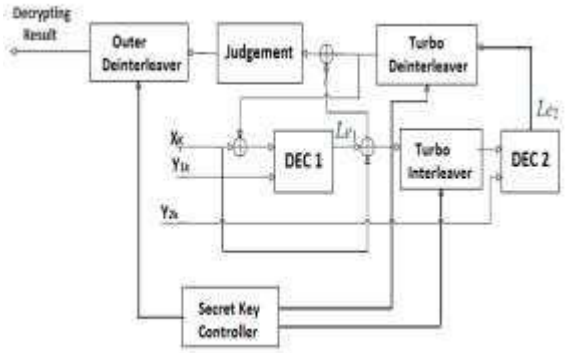
Sources: Authors Compilation.

The *d denotes* standard deviation *that* determines the number of 0s in the sequence *P*, the table shows the relation between the standard deviation and coding rate a high standard deviation means better security of the encryption scheme. If *d* equals to 0, the dynamic puncturing mechanism becomes periodic puncturing scheme of the normal Turbo code, the coding rate is 1/2, and the information can not be encrypted. Therefore, in order to provide high security require high coding rate suitable value of *d*.

B. Decryption Process

During the decryption, the receiver firstly generates the sequence *P* using the secret key *s*, and classifyX_k, Y_{1k}, and Y_{2k} in the received sequence according to *P*. Then the receiver sends them to a Turbo decoder, shown as Fig. 5.Among the three kinds of received bits, X_k and Y_{1k} are input to the first decoder (DEC1), and X_{ik} and Y_{2k} are input to the second decoder (DEC2). After that, Turbo decoder will start iterative decoding process.

Figure-5: Block Diagram for Decryption Process



Sources: Authors Compilation.

In the Turbo decoding process, the Log-MAP algorithm is implemented In Log-MAP decoding algorithm, the decoder decides dˆ k or otherwise compute the Logarithm of Like lihood Ratio. The soft output from each constituent decoder is divided into three parts: the extrinsic output which is new information derived by current stage of decoding ,a weighted version of the systematic input , and a copy of the input a priori information The Turbo decoder judges the result ˆdˆ k according to L (d_k) after several iterations. In the decoding process only a exact user can generate the right P, which is equal to that of the transmitter. So, sustainable information encryption and decryption can be achieved by this way.

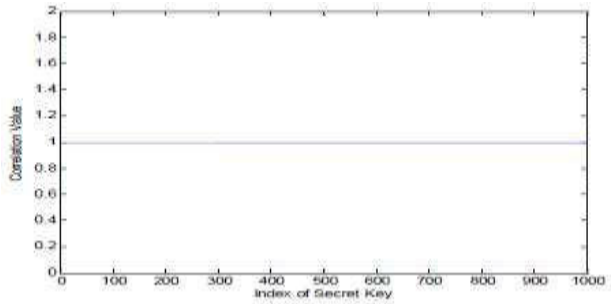
Results and Discussions

The section will present the simulation environment occupied to evaluate the proposed scheme and present simulation results that shows the high level of the security to the proposed encryption scheme. Simulation is carried out using MatLab software here the numbers ranging from 0 to 255 are taken as input data the channel platform is the AWGN. Generate Matrix (g) for this simulation is [1 1 1;1 0 1] with 5 as Iteration number .The Frame Size is 400 bits and for Decoding uses the Log-MAP Algorithm.

A. Security based on Proposed Encryption Method

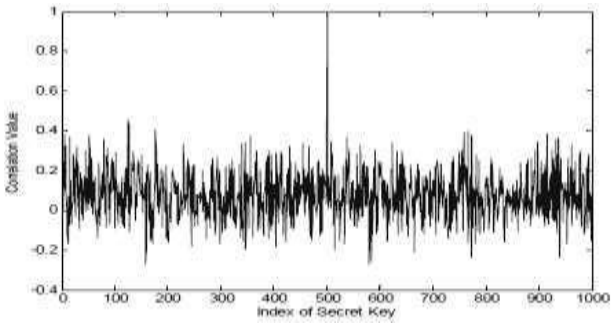
If the attacker will try to decode encrypted data by using the wrong key they obtain the incorrect results so our encryption method is obviously high secured. Here the correlation value is used to measure the similarity between the decrypted and original data. In the following section, to simulate the attack scenarios 1000 randomly generated secret keys are used in which only the 500th is correct one. Below shows the correlation results between original input data and the decoded data using different coding rate and keys. The coding rate is ½ this indicates that the standard deviation of that sequence is zero. In fig.6 shows that original data is same as the decrypted data. This means the coding scheme cannot encrypt information.

Figure-6: Correlation Result (Code Rate=0.5)



Sources: Authors Compilation.

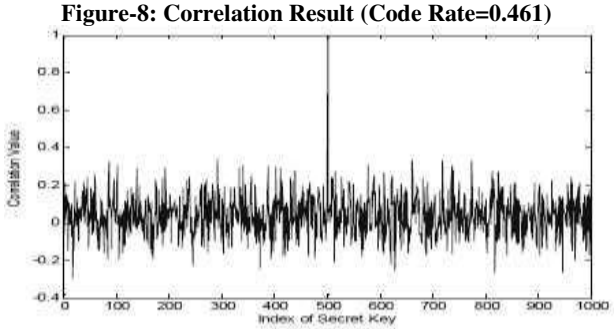
Figure-7: Correlation Results (Code Rate=0.482)



Sources: Authors Compilation.

Figure7 and 8 shows results of the proposed encryption method. These results indicate the reliable security when a proper coding rate is selected. With the increase of the standard deviation of the sequence, the difference of the decrypted data with correct and the wrong keys becomes high. And there is only one correlation peak in the location of the 500th key others have low correlation values.

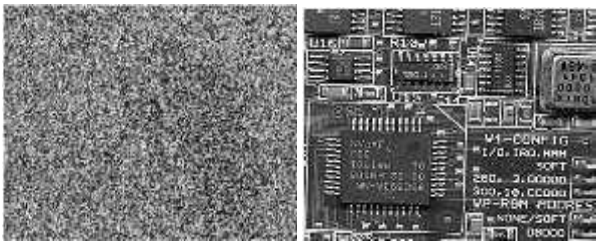
Figure-8: Correlation Result (Code Rate=0.461)



Sources: Authors Compilation.

Fig. 9 and 10 indicates the results of experiment which demonstrate the applications of encryption scheme based on Turbo-codes in image encryption. Picture (a), and (c) represents the original images then (b) and (d) represents the image that decrypted using a wrong key with 0.3 as the standard deviation of the sequence. All these results which indicate that our Turbo based encryption scheme gives effective image encryption.

Figure-9: Simulation Results of a PCB Circuit



a) Original Image **b) Encrypted Image**

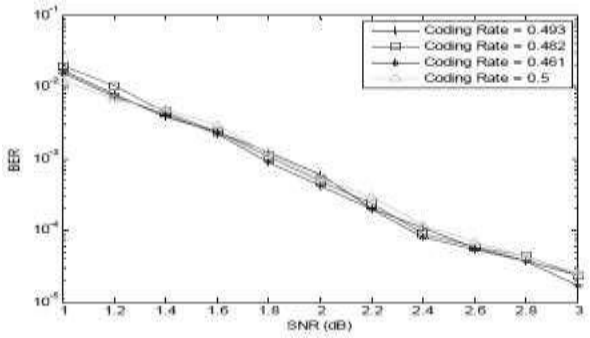
Sources: Authors Compilation.

B. BER Performance for the Proposed Method

The bit error rate of the encryption method is more relevant one based on the error-correcting code. Fig.10 shows BER performances with different coding rates for proposed encryption scheme. From the detailed analysis of curve gives us the finding that BER performance does not vary with small variation in the coding rate for our propose decryption scheme.

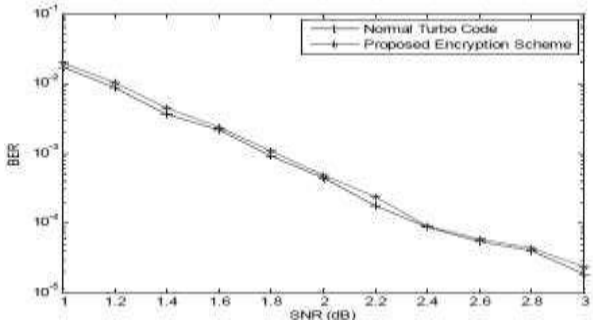
Figure-10: BER Value with Various Coding Rate for Proposed Scheme

Sources: Authors Compilation.



The Fig. 11 shows the BER performances of the proposed encryption code and normal Turbo code with same coding rate. From the detailed analysis of this figure when the SNR of the AWGN channel varies from 1 dB to 3dB the coding rate in this experiment is 0.482. BER performance of proposed encryption scheme is as excellent as that of the normal Turbo code, so the dynamic puncturing mechanism does not decrease the error correction capability of the Turbo code.

Figure-11: BER Curve for Proposed Scheme with Normal Turbo Code



Sources: Authors Compilation.

CONCLUSIONS

Turbo codes are the advanced error correction code which is also applicable with the encryption scheme enables high rate of security with error immunity. The information encryption is achieved by means of a key-controlled dynamic puncturing mechanism.

According to the simulation result shows that BER value of the proposed system is very low and correlation results indicates high level security. The application of our proposed method in image encryption also an apt one. Hence encryption scheme based on turbo code using dynamic puncturing mechanism is the efficient and reliable one for the different communication system.

REFERENCES

[1] Claude, Berrou; Ramesh, Pyndiah; Patrick, Adde; Catherine, Douillard, and Raphael, Le Bidan, (2005). "An Overview of Turbo Codes and their Applications", The European Conference on Wireless Technology.

[2] Branka, Vucetic; Yonghui, Li; Lance, C. Perez, and Fan, Jiang, (2007). "Recent Advances in Turbo Code Design and Theory", Proceedings of IEEE, June, Volume: 95, Number: 6.

[3] C., Berrou; A., Glavieux, and P., Thitimajshima, (1993). "Near Shannon Limit Error-Correcting Coding: Turbo Codes", Proceedings of IEEE International Conference on Communications, ICC '93, Geneva, May, pp. 1064–1070.

[4] W., Stallings, (2004). "Cryptography & Network Security: Principles and Practice", 3rd ed., Prentice Hall, August.

[5] M., C. Valenti, and J., Sun, (2004). "Turbo Codes", Chapter 12 of Handbook of RF and Wireless Technologies, (Editor: F. Dowla), Newnes.

[6] Burr, A.; (2001). "Turbo-Codes: The Ultimate Error Control Codes?", IEEE Electronics & Communication Engineering Journal, August, Volume: 13.

[7] J., Hokfelt; O., Edfors, and T., Maseng, (2001). "A Turbo Code Interleaver Design Criterion Based on Performance of Iterative Decoding", IEEE Communications Letters, Volume: 5, Number: 2.

[8] S., A. Barbulescu, and S., S. Pietrobon, (1999). "Turbo Codes: A Tutorial on a New Class of Powerful Error Correcting Coding Schemes, Part 2: Decoder Design and Performance", Journal Elec. and Electron. Eng., Australia, September, Volume: 19, pp. 143-152.

*****

SKR TREE FOR SPATIO – TEMPORAL QUERIES BASED ON CIRCULAR SAFE ZONE

Saranraj G.[114] Mookhambika N. [115]

ABSTRACT

Moving object database (MOD) engine is the foundation of Location Based Service (LBS) information systems. Continuous queries are important in spatial-temporal reasoning of a MOD. The communication costs were the bottleneck for improving query efficiency until rectangular safe region algorithm partly solved this problem.

However, this algorithm can be further improved, as we demonstrate with the dynamic interval based continuous queries algorithm on moving objects. Two components, circular safe region and dynamic intervals were adopted by our algorithm. Theoretical proof and experimental results show that our algorithm substantially outperforms the traditional periodic monitoring and rectangular safe region algorithm in terms of monitoring accuracy, reducing communication costs and server CPU time. Moreover, in our algorithm, the mobile terminals do not need to have any computational ability.

KEYWORDS

RFID, GPS, Wi-Fi, Location, Service, Moving Object Database, Continuous Spatial Query etc

INTRODUCTION

The development of technology has made it possible to track moving objects such as vehicles, aircrafts, vessels, wildlife, and human objects such as firefighters in a fire field. Technologies such as global positioning system (GPS), radio-frequency identification (RFID), cellular wireless networks (such as commercial cellular phone networks) and even triangulated wireless fidelity (Wi-Fi) networks can all provide location information in real- time, although at different precisions with different ef- fective ranges.

Two major trends can be identified to manage the large amount of location and property information that varies with time: moving object databases (MOD) and data stream technology (DST). The first approach im- plies extending traditional database techniques with models and index structures suitable to track the loca- tions of the moving objects efficiently. The second ap- proach focuses on the processing of continuous location updates as they arrive. The boundary between these two approaches is not always clear in relation to the topic of this survey: Both propose alternatives to classical data- base techniques, which are not considered appropriate to manage the continuously changing locations of the moving objects [1]. Our research will focus on the MOD ap- proach but can be easily modified to adapt to the DST approach as well since our Dynamic Interval Based Cir- cular Safe Region (DIBCSR) algorithm requires the minimum frequency of location updates which can be provided by both approaches.

[114] PG Scholar, Dhanlakshmi Srinivasan College of Engineering, Tamil Nadu, India, saran4friends@gmail.com
[115] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mookhambika@gmail.com

MOD is a system that performs storage management and query analysis on time-variable spatial information of moving objects [2-4] which combines multiple disci- plines and research areas including geographical infor- mation systems (GIS), spatial databases, spatial-temporal databases, computer graphics, computational geometry, artificial intelligence and mobile computing.

Application of MOD requires the optimal efficiency of the queries which can only be provided by continuous spatial-temporal queries. A regular spatial-temporal query only returns a single result set. In contrast, a continuous spatial-temporal query returns result sets continuously from the registration to the cancellation of the query, which is called the effective period of the query. Even if the query conditions remain unchanged during the effect- tive period, the query result must be updated continu- ously due to the continuous movement of the queried objects. Here are two examples of continuous spatial- temporal queries which provide commonly used LBS such as range query or the k-nearest neighbor (kNN) query:

1) List all vehicles that appear in region R in the next 10 minutes.
2) Continuously mark the ten closest vehicles to gas station number five.

These types of queries are not commonly supported by traditional relational database engines. In order to facili- tate these continuous spatial-temporal queries, a MOD engine must implement the query processing and ideally, with optimal performance at a low cost.

RELATED WORKS

Performance of the dynamic updates of the query result set during the effective period is the main research topic of MOD and spatial-temporal reasoning. In order to per- form continuous query optimization in a distributed sys- tem, not only the query cost must be minimized, the communication cost for updating location information of the terminal devices must also be minimized. However, most of the previous works on continuous queries have focused on reducing the query cost and has ignored the communication cost [5-9], in which the occasions for reporting location information are determined by the terminal device at fixed intervals or when the object's location (constant distance interval) experiences a significant change. This class of uniform time / distance in- terval strategies has the following weaknesses:

- The location updates are not adaptive to queries. When queries are scarce or there are no queries at all, a large amount of communication bandwidth and battery power of the terminal device may be wasted on the up- dates.
- Low efficiency of queries could cause inconsistency with reality. In the periodical update strategy, improving the consistency of the query result with reality relies on the location update frequency increase. This means higher communication costs and may even makes improvement impossible because of the bandwidth and network delay limitation.

- Unbalanced workload is applied on the server. In order to improve the reality consistency, the server must update large amounts of location information constantly and recalculate all the queries. An overloaded server usually means low responsiveness and poor reliability.

Hu, et al. [10] proposed a continuous query update strategy based on a rectangular safe region (RSR) me- thod which can alleviate the previous three problems. However, with analysis and experiments, we found that this strategy requires considerable computation power on terminal devices. This performance bottleneck may be- come more significant with a larger query load.

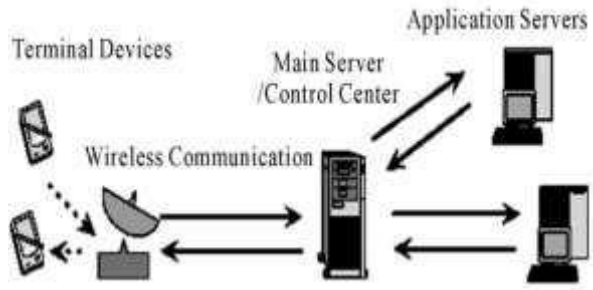
We propose an improved dynamic interval based cir- cular safe region (DIBCSR) strategy to replace the RSR strategy. In our strategy, the location update information is updated dynamically by the server and the terminal devices do not need to download the safe region infor- mation. Experiments show that our strategy eliminates the computation requirements of the terminal devices with equal or better system performance than RSR stra- tegy.

Moreover, most of the previous studies only support one specific query type, such as either range queries kNN queries but not both. Our proposed DIBCSR algorithm supports both range queries and kNN queries.

SAFE REGION BASED LOCATION UPDATES

Figure 1 demonstrates the infrastructure of a moving object query system. The kernel of the system is the con- trol center (the main server of the system) in the center of the figure which runs the MOD engine, collects location information, handles continues queries and provides query results to the application servers to the right of the figure. Therefore, the major computation workload is applied to the main server/control center of a MOD system. For simplicity, we refer to the main server/control center as server in this paper.

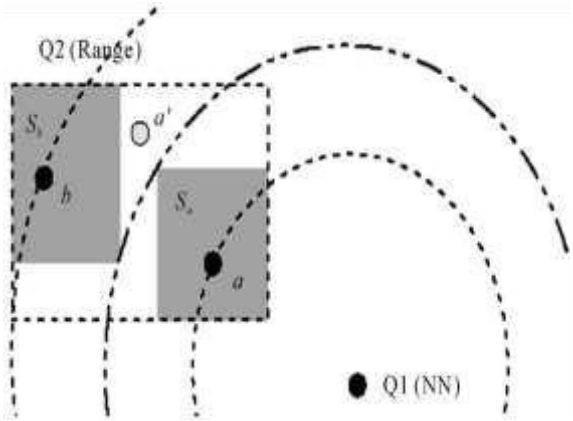
Figure-1: Infrastructure of MOD System



Sources: Authors Compilation.

Terminal devices, which are the monitored moving objects, obtain their own location information from the GPS system and transmit it to the server via a wireless communication network. The whole system's timeliness and efficiency is affected by the wireless communication bandwidth. The location information updates are often the bottleneck because of the limited wireless bandwidth and the high sampling rate in the traditional uniform time/distance interval strategies.
The idea behind the rectangular safe region (RSR) algorithm [10] is to define a rectangular safe region for every object according to the registered query and the latest location obtained. As long as the object's motions do not exceed its safe region, all the query result sets of the object remain unchanged (Figure 2). The terminal device is informed of the safe region assignments dy- namically. Hence when a terminal device finds that it has exceeded the safe region, it will report its new location information. E.g., when an object a in Figure 2 has moved out of its safe region of Sa to location a', it will report its new location to the server which will recalcu- late the results of a continuous k-nearest-neighbor (kNN) query Q1 and a range query Q2.

Figure-2: Rectangular Safe Region



Sources: Authors Compilation.

Through analysis and experiments, we found that although the RSR algorithm is effective, it has the follow ing weaknesses which can be improved:

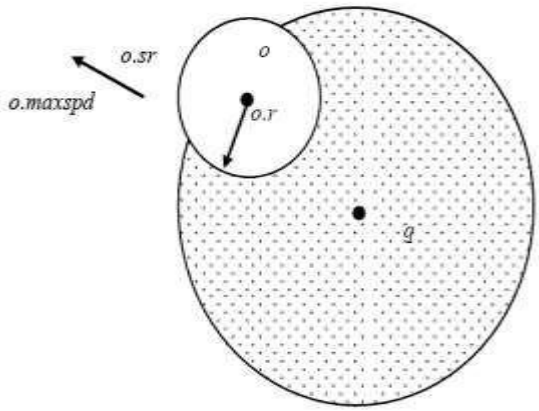
1) RSR requires that the terminal device has memory to store its current safe region and computing power to determine whether it has exceeded the safe region. However, in practice, many low-cost terminal devices (e.g. a GPS dog collar) do not have memory and computing power in addition to GPS satellites communication.

2) In RSR, data communication is bidirectional. The terminal devices not only need to upload location infor- mation to the server, but also need to download safe re- gion information from the server. When the query fre- quency increases, the frequency of safe region download to the terminal devices increases. When the query fre- quency is high enough, the communication cost may be even worse than the uniform time interval (UTI) strategy. We have a detailed analysis of this problem in Section 5.2.

3) Computations involved in the RSR strategy are complicated, especially for kNN queries.

On the observation of these problems, we propose a new continuous query algorithm. We define the safe region of object *o* (referred to as *o.sr*) as a circle with the center at the location of the object and the radius of *o.r* (Figure 3). Assume the maximum speed of the object is *o.maxspd,* then continuous query result of query *q* will not be affected within the time interval of *o.r/o. maxspd*. Hence the server can issue a location report query to the terminal device at the time of *(o.r/o.maxspd - δ)* where δ is the sum of communication and computation delays.

Figure-3: Location Update of a Moving Object with Circular Safe Region under Continuous Query



Sources: Authors Compilation.

In comparison, the advantages of our DIBCSR strategy are:

1) The terminal device does not need to have any computing power. The only task of the terminal device is to report its location upon the request of the server. This is determined by our main algorithm which is given in Sec- tion 4.1.

Moreover, the location update sampling re- quests are distributed by the server. Therefore, when the sampling strategy needs updating, such as when safe regions are reassigned because of objects' movements, only the server is affected and the communication cost will not be affected.

2) The algorithm determining the assignment of a cir- cular safe region is simpler than a rectangular one. There- fore the computation is reduced for safe region assign- ments. We provide the detailed safe region assignment algorithms for continuous range queries and continuous kNN queries in Sections 4.2 and 4.3 respectively.

3) The updates of safe regions have been minimized with the selection of circular shape safe regions and therefore the communication cost is minimized. We pro- vide mathematical analysis in Section 5.1.

4) The communication cost is reduced in comparison with the RSR strategy. Detailed analysis of this feature is provided in Section 5.2.

5) Computations involved in RSR strategy are complicated, especially for kNN queries. In contrast, computations are much more concise in our DIBCSR strategy.

DYNAMIC INTERVAL BASED LOCATION UPDATES

We use an C++/Java style pseudo code syntax, including comment syntax of double slash, to represent the algo- rithms in a more concisely and precisely. Properties of the moving object *o* and continuous query *q* are ex- plained in **Table 1** and **Table 2**.

A separate process will be responsible for determination of the objects that are due for reporting new locations and sending the requests. The main algorithm which runs on the server is as following:

Table-1: Properties of a Moving Object

q.region	Query region of a range query
q.result	Query result set
q.effUTI	the query effective period
o.circle.p	the center of the circular query region
o.circle.r	the radius of the circular query region

Table-2: Properties of Continuous Query

o.p	Last reported object location
o.r	Radius of the object safe region
o.sr	Object safe region
o.maxspd	Maximum speed of the moving object
o.upt	Next location update time of the moving object

Algorithm*-1: *Main Algorithm for Continuous Query Processing

```
//OList is the object list, QList is the query list while
(received query q and
current time t within q.effUTI)
do
{
if (q is newly registered) then
{
//new a query q for processing, either //range or kNN
query
NewQuery(q);
}
if (q is cancelled) then remove q from QList;
if (q is location update of object o) then
{
//update safe region of object o and //related query result sets
UpdateSR(QList, o);
o.upt = t + o.r/o.maxspd - delay;
}
}
```

```
UpdateSR(QList, o)
{
//range query location updates processing Update-
SRA(QList, o);
//kNN query location updates processing Update-
RSR(QList, o);
}
```

4.1. The Main Algorithm for Continuous Query Processing

Both continuous range query and continuous kNN query are supported by our strategy. A continuous range query is one that returns all the objects in q.region within the query effective period where query region can be either rectangular or circular. A continuous kNN query is one that returns the closest k objects to the query location. An ordered kNN query requires the results to be returned in increasing order and an unordered kNN query does not request the results to be in order. An ordered kNN query is what we will consider in this paper and which is more complicated than an unordered kNN query. These two different types of continuous queries require different new query processing and location update processing algorithms which we present in different sections as follows:

4.2. Continuous Range Queries

The query processing and location update algorithm for continuous range queries are as follows:

Algorithm-2: New Range Query Processing

```
NewQuery(q)
//q is a range query
{
for (∀o ∈ OList) do
{
qResult = Req(q, o.sr);          //query result
if (qResult = "Y" ) then
q.result= q.result ∪ {o};
// added to the result set of q
// if the query result of is undecided
if (qResult = "U" ) then
{
Query location of object o;
//object o must report immediately
//recalculation is then performed.
UpdateSR(QList, o);
}
}
}
```

In a range query, the Req function in Algorithm 2 which is the query result of safe region is defined as:

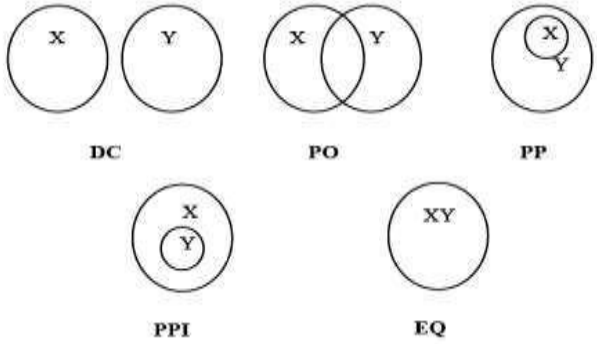
$$\text{Req}(q, o.sr) = \begin{cases} Y & \text{if } RCC5(q.r, o.sr) = PPI \\ N & \text{if } RCC5(q.r, o.sr) = DC \\ U & \text{otherwise} \end{cases} \quad (1)$$

A return value of "*Y*" indicates that the safe region is inside the query region and therefore object *o* is within the result set. A return value of "*N*" indicates that the safe region is outside of the query region and therefore the object *o* is not included in the result set. A return value of "*U*" indicates that the safe region intersects with the query region. The result is therefore undecided. Hence the precise location of the object needs to be obtained in order to recalculate. The region connection calculus (*RCC*) serves for qualitative spatial representation and reasoning and *RCC*-5 is a widely used binary rela- tionship model in automated spatial reasoning [11] with five binary relationships {DC, PO, PP, EQ, PPI} (dis-creteness, proper overlap, proper part, equivalence, proper inclusion) demonstrated in **Figure 4**. The function *RCC*5(X, Y) returns the *RCC*-5 relationship of X and Y for topology analysis.

Algorithm-3: Range Query Update

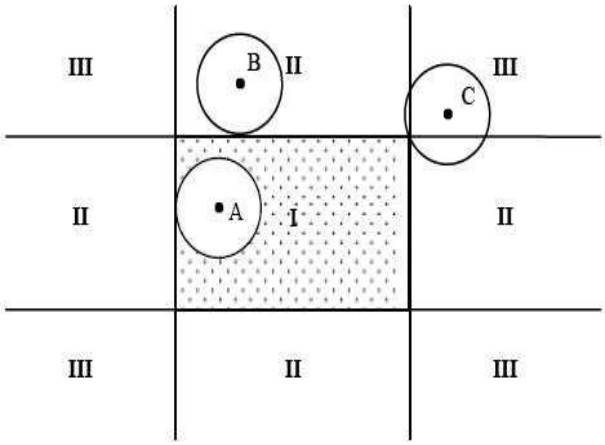
```
//The functionality is to update the range query
//result set and check/update the object safe //region by invoking
the SafeRegion sub //function.
UpdateSRA(QList, o)
{
    o.r = ⊓;
    for (¯q ⊓ QList and q is a range query) do
    {
    // If safe region function returns r>0
    // then o is within the result set of q
    if SafeRegion(q, o, r) then{
    q.result = q.result ⊓ {o};
    return true;
    }
    }
    }
```

Figure-4: Binary Spatial Relationships in RCC-5



Sources: Authors Compilation.

Figure-5: Rectangular Range Query Safe Region.



Sources: Authors Compilation.

Algorithm-4: The Calculation Of Safe Region For Range Query

The purpose of this algorithm is to calculate the safe region radius and decide the query result set. The safe region radius *r* is returned for the object *o* under query *q*. The Boolean return value of *true* or *false* indicates wheth- er object *o* is within the query result set.

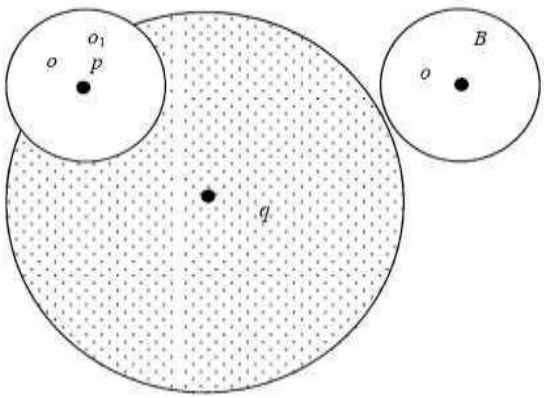
```
SafeRegion(q, o, & r)
{
if (q is a rectangular range query) then
{
//Three circumstances exist,
//as demonstrated in Figure 5.
//A: o.p is inside query region I
//B: o.p is inside query region II
```

```
//C: o.p is inside query region III
if (o.p is inside query region I) then
{
r = distance from o.p to the closest edge of the rectangular query region;
return true;
}
else if (o.p is inside query region II) then
r = distance from o.p to the closest edge of the rectangular query region;
else // o.p is inside query region III
r = distance from o.p to the closest vertex of the
        rectangular query region;
return false;
}
else if (q is a circular range query) then
{
//Two circumstances exist,
//as demonstrated in Figure 6.
//A: object o is inside the
//circular query region
//B: object o is outside of the
//circular query region
        doq= dist(o.p, q.circle.p);
//dist(a, b) represents the distance
//between point a and point b. doq is the //distance between object o and query q.
r = ABS(doq - q.circle.r);
if (doq <= q.circle.r) then
return true;
else
return false;
}
}
```

Theorem-1: When the motion of the object in Algorithm 4 does not exceed the safe region, the result set of the continuous range query does not change.

Proof: *Apparently, under the circumstance, o.sr does not intersect with q.r. Hence object o's motion inside o.sr will not affect the query result set.*

Figure-6: Circular Range Query Safe Region



Sources: Authors Compilation.

4.3. Continuous kNN Query

Algorithm-* 5: *An Ordered kNN Query Processing

```
NewQuery(q)

//q is an ordered kNN query
{
1) Decide the object list OList near query location q.p based on the spatial-temporal index of moving objects;
/*e.g. objects within neighboring rectangles can be selected in an R*-tree indexed system. This step reduces the object set that is processed to reduce the following computation.*/
2) Perform sorting to the objects in OList by dist (o.p, q.circle.p) ascending, the nearest k + 1objects are saved in q.result;
/*Quick Sort algorithm is applied and the Compare function is given below. The reason why we save the (k +1)th object is for the calculation of the safe region.*/
3) Update the safe regions for all objects;
}
```

Algorithm-*6: *Distance Comparison Algorithm for Objeects

```
For simplicity, we use q to represent q.circle.p and object names o1, o2 to represent the object location o1.p and o2.p in this section.
// Function returns –1 when o1 is nearer than o2; // returns 1 when o1 is farther than o2.
// Since all calculations are floating-point, we do // not consider the equal scenario.
Compare (q, o1, o2)
{
if ( dist(q, o1) + o1.r < dist(q, o2) – o2.r ) then
   return –1;
if ( dist(q, o1) – o1.r > dist(q, o2) + o2.r ) then
   return 1;
Query locations of o1, o2;
//Distances of safe regions to q overlaps, //query precise locations for further //comparison.
o1.r = 0;
o2.r = 0;
if (dist(q, o1) < dist(q, o2) ) then
   return –1;
else
   return 1;
}
```

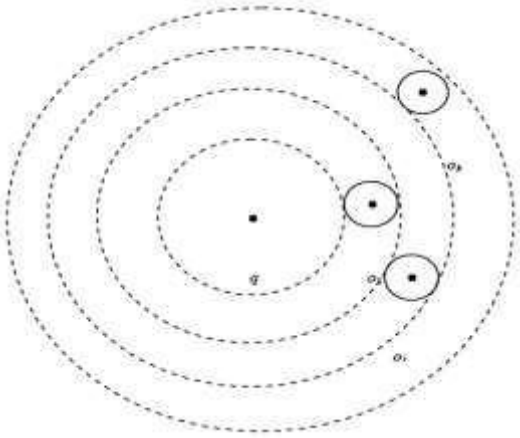
4.4. Circular Safe Region Calculation and Updates for Continuous kNN Queries

Following is the formula to update the safe region radius for the i^{th} object in the object set ascending sorted by distance to ordered kNN query q in Algorithm 5. **Figure** 7 shows an example of safe regions assignment in such an ordered kNN query q. The first object in the result set is q and the extra object o_{k+1} is kept for calculation of safe region radius of o_k.

$$o_i.r = \begin{cases} \text{if } 0 < i \le k, \text{ in the result set, then} \\ \quad \left[\min\left\{ o_i.r, \frac{dist(o_i, o_{i-1})}{2}, \frac{dist(o_i, o_{i+1})}{2} \right\} \right] \\ \text{if } i > k, \text{ out of the result set, then} \\ \quad \min\{o_i.r, dist(o_i, q) - quar(q)\} \end{cases} \quad (2)$$

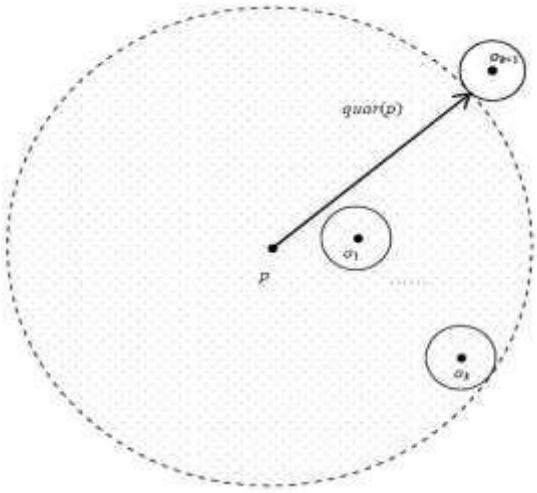
Where $quar(q)$ is the radius of the quarantine region for query q which surround and only surround the safe re- gions of all objects in the result set. Therefore, $quar(q) = dist(o_k, q) + o_k$. **Figure 8** shows how such a quarantine region is assigned for such an ordered kNN query q. Hence we have the following property: either in or out of the kNN query result set (inside or outside the quarantine region), none of the safe regions of objects overlaps with each other. Therefore, when all the objects are moving inside their own safe regions, the result set and its order are not affected.

Figure-7: Safe Regions Assignment in an Ordered kNN Query



Sources: Authors Compilation.

Figure-8: Quarantine Region Assignment for an Ordered kNN Query



Sources: Authors Compilation.

4.5. Location Update Processing for Continuous kNN Queries

For continuous kNN query, when object location is up- dated, one of the four following scenarios will happen. The detailed update algorithm is given in Algorithm 7:

1) Original location was inside the quarantine region and new location is also inside the quarantine region: order adjustment in the result set is necessary.
2) Original location was outside the quarantine region but new location is inside the quarantine region: a new query is necessary to recalculate the complete result set.

3) Original location was inside the quarantine region but new location is outside the quarantine region: same as 2).

Both original location and new location are outside the quarantine region: only need to update the object's safe region.

Algorithm-7: Location Update Processingfor kNN query

//Purpose is to update safe region and result set
UpdateRSR(*QList*, *o*)
{
for ($\forall q \in QList$ and q is a kNN query) do
{
if (($o \notin q.result$ and $dist(o, q) \le quar(q)$) or ($o \in q.result$ and $dist(o,q) > quar(q)$)) then
{
Execute Algorithm 5;
//Processed as a new
query continue;
}
else if ($o \in q.result$
and $dist(o, q) \le quar(q)$) then
{
//Original result set of q is $\{o_i | i = 1,\cdots, k\}$, //if object o's original index is i and //index after the new sorting by $dist(q, o)$ //is i'
Execute Algorithm 5 within the range of

$$\begin{cases} i-1, i, \cdots, i', i'+1, & \text{when } i' \ge i \\ i'-1, i', \cdots, i, i+1, & \text{when } i \ge i' \end{cases}$$

Use result set from Algorithm 5 to replace the respec- tive subset in the original result set;
continue;
}
else //$o \notin q.result$ and $dist(o, q) > quar(q)$
{
Adjust $o.r$ following the Req function given by Equa- tion (1);
continue;
}
}
}

ANALYSIS AND EXPERIMENT

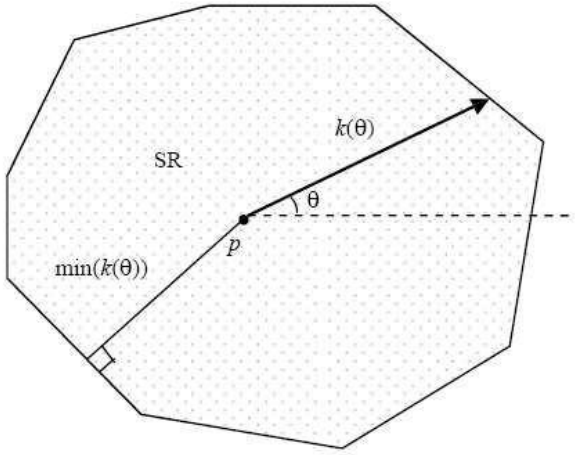
5.1. Analysis of the Safe Region Shape

As previously mentioned in Section 3, one reason for selecting the circular safe region shape is to minimize the updates of safe regions and therefore the associated communication cost. We further provide mathematical analysis here.

In a safe region based strategy, the communication cost of a location update is inversely proportional to the minimum location update time. This is because the ob- ject's motion direction is generally unpredictable. There- fore its probability of leaving the safe region through any portion of the border is equal. Assume *SR* to be the safe region, *p* to be the last reported location. If in the direc- tion of θ, the distance from *p* to the edge of the safe re- gion is *k*(θ) (**Figure 9**), then the minimum location up- date time is:

$$1/cost_{communication} \propto time_{update} = \min\left(k(\theta)\right)/o.maxspd \quad (3)$$

Figure-9: Shape of Safe Region.



Sources: Authors Compilation.

Assuming *o.maxspd* is a property of the object that we cannot control, our goal is to maximize the min(*k*(θ)) in order to maximize the t_{update}. When a specific shape of safe region is selected, increasing the area of the safe region is obviously going to increase min(*k*(θ)). However, in a range query or a kNN query, the largest area is eventually bounded by the objects' distribution and the size of the query region. Therefore, if we assume the area of the safe region is determined, then the maximized average distance from *p* to the edge of the safe region in all directions, min(*k*(θ)), will come with the isotropic safe region – a circle.

5.2 Analysis of the Safe Region's

Communication Cost

In the RSR strategy [10], data communication is bi-di- rectional: the terminal devices upload location informa- tion and download the safe region information. However, the authors only discussed the communication cost of location information upload, which is not precise. We take the bi- directional data communication into consid- eration in the following discussion and then compare the communication cost with our DIBSCR strategy. In the RSR strategy, communication in the down-link direction from the server to the terminal device transmits information of the rectangular safe region, each deter- mined by two points or four coordinates. Communication in the up- link direction from the terminal device to the server transmits a location update, each includes one point or two coordinates. In newer wireless networks such as Wi-Fi, Wi-Max or 3G, data is transmitted in data frames (called synchronous transmissions mode). Since the data amount to be transmitted/received by the ter- minal device every time is quite small which can easily fit into a single frame, the location update rate is the main factor affecting the communication cost. This loca- tion update rate is constant for UTI strategy while varia- ble for safe region based strategies. Assume in the RSR strategy, the safe region update rate is *s* which depends on the query rate of *q*. We therefore represent it using the function of *s*(*q*) which increases with the query rate of *q*. And assume the location update rate is *u* and the com- munication delay is *d*. The total communication cost is (*s*(*q*) + *u*)·*d* which increases with the query rate. Hence when the query rate is high, the RSR strategy can be even worse than the UTI which makes it not feasible.

In contrast, in our DIBSCR strategy, terminal de- vice does not need to download safe region information which only leaves the location update term of *u*·*d*. Moreover, the location update rate *r* is most equal to the constant rate in the UTI strategy because it is based on the dynamic time interval which is greater than or equal to a preset value. We further provide the estimated loca- tion update rate *r* in DIBSCR as following:

The basic estimate of the location update interval is derived from Equation (3):

$$1/u \propto t_{update} = o.r/o.maxspd \quad (4)$$

where, t_{update} is the minimum amount of time the object may exceed the circular safe region and therefore re- quests a location update. The estimate of the location update rate *u* is therefore relying on the estimate of the object's maximum speed *o.maxspd*.The estimate of *o.maxspd* can be either fixed (for in- stance the object types of pedestrian, motor vehicle or high-speed train) or it could also be further refined by prediction from the object's historical locations.

This could reduce the communication cost when the object is temporarily immobile (such as when the pedestrian stopped by a coffee shop or when the motor vehicle is parked). Certainly in any prediction based speed estimate, we need to be on the conservative side and control the computation cost although there are outstanding predic- tion methods such as Back Propagation Networks (BPN). For simplicity, we do not want to include consideration of a missing rate. One possible conservative estimate is:

$$o.maxspd = \min\left(fixed\_max\_speed,\ v_{last} + max\_acceleration * (t_{current} - t_{last})\right)$$

where

fixed_max_speed: the maximum possible speed for an object type

v_{last}: calculated speed at the last location report time

max_acceleration: the maximum possible acceleration of an object type

$t_{current}$: the current system time

t_{last}: the last location report time.

Either the *o.maxspd* is fixed or further bounded by a prediction, it is tightly bounded and hence the minimum amount of time the object may exceed the circular safe region and the maximum location updated rate is tightly bounded and independent of the query rate.

5.3. Experimental Analysis of the System Performance

In order to further evaluate our strategy, we constructed a simulation system to evaluate the UTI strategy, RSR strategy and our DIBCSR strategy. In our simulation system, object *o*'s motion direction and speed are ran- domly generated. The object's speed should not exceed a fixed maximum speed of *o.maxspd*. In UTI simulation, we have two location update intervals of 0.1s and 1s, referred to as UTI(0.1) and UTI(1) respectively in the simulation results.

In our experimental analysis through simulation, three comparison criteria are applied: 1) precision, 2) commu- nication cost and 3) server workload. We analyze the si- mulation results separately in the following section.

1) Precision

The precision of a continuous query result is defined as: at time *t*, the system query result is RESULT(*t*); the actual object set that satisfies the query condition is TRESULT(*t*), standing for the true result. In order to use a higher value to represent higher precision, we define the *equal*(*x*, *y*) function to return 1 when $x=y$ and 0 when $x \neq y$. In the time interval of [a, b], we average the equality between the returned value and the true value, and define precision as:

$$\frac{1}{b-a}\int_b^a equal\left(RESULT(t), TRESULT(t)\right)\mathrm{d}t$$

The precision is obviously affected by the communi- cation delay since it causes the difference between the actual location and the reported location. The result is represented in Figure 10 and the precision of DIBCSR and SBR are approximately the same and both are better than UTI.

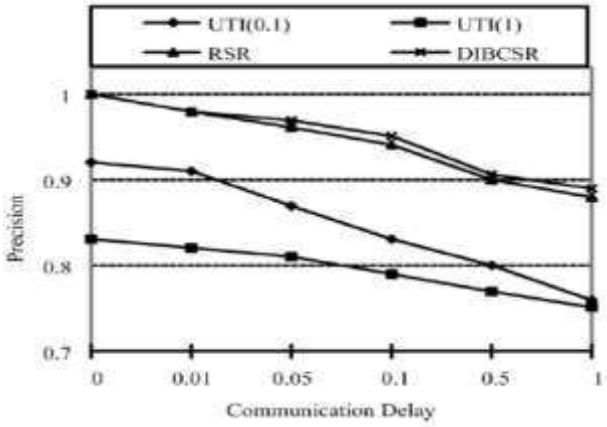
The simulation results shown in Figure 11 confirm our analysis in Section 5.2 that the performance of DIBCSR is significantly better (lower communication cost) than RSR when considering bi-directional communication cost; and at high query rate, communication cost of RSR can be even worse than UTI with a larger time interval (lower sampling rate).

2) Communication Cost

3) System Scalability

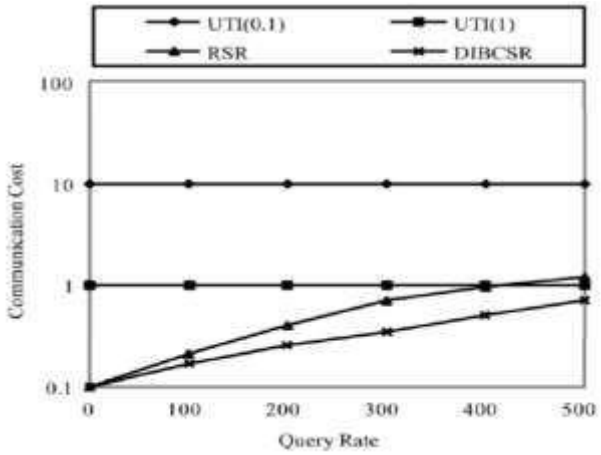
Figure 12 shows the comparison of server workload for different strategies under the query rate increase. Since the workload is balanced better, both DIBCSR and RSR apply less workload on the server than UTI. Because we further simplified the safe region computation by applying circular safe region, DIBCSR applies even less workload than RSR on the server. This advantage is more significant under the query rate increase. This low server workload feature of DIBCSR helps to improve the system scalability.

Figure-10: Comparison of Precision



Sources: Authors Compilation.

Figure-11: Comparison of Communication Cost



Sources: Authors Compilation.

CONCLUSIONS

This paper analyzes the weaknesses of the RSR strategy and proposes a DIBCSR strategy to replace the RSR strategy for continuous queries in MOD. Theoretical analysis and simulation experiment both show that the new strategy has multiple advantages. Firstly, the new strategy does not require computation over the terminal devices.

Therefore, cost of the terminal devices is reduced under the precondition of equal or better system performance. Secondly, terminal devices do not need to download the safe region information from the server which reduces the communication cost effectively. Finally, computation is simplified by applying circular safe regions.

Hence the server workload is reduced which improves the system scalability. Possible future works of the research include implementation of the strategy in an applied MOD engine for

an information system providing LBS to public transportation, taxis and private vehicle devices or pedestrians with hand-held mobile devices. Application of our strategy potentially provides real-time range queries and kNN queries to support LBS at a low cost with a high performance in addition to system de- sign and implementation ease and flexibility.

REFERENCES

[1] S., Ilarri; E., Mena, and A., Illarramen di, (2010). "Location-Dependent Query Processing: Where We Are and Where We Are Heading", ACM Computing Surveys, Volume: 42, Number: 3, pp. 1-79, doi:10.1145/1670679.1670682.

[2] H., D. Chon; D., Agrawal, and A., El Abbadi, (2001). "Storage and Retrieval of Moving Objects", Proceedings of 2nd International Conference on Mobile Data Management, Hong Kong, 8-10 January, pp. 173-184.

[3] H., D. Chon; D., Agrawal, and A., El Abbadi, (2003). "FATES: Finding a Time Dependent Shortest Path", Proceedings of 4th International Conference on Mobile Data Management, Melbourne, 21-24 January, pp. 165-180.

[4] Y., Chen; F., Rao; X. Yu, and D., Liu, (2003). "CAMEL: A Moving Object Database Approach for Intelligent Location Aware Services", Lecture Notes in Computer Science, Volume: 2574, pp. 331-334, doi:10.1007/3-540-36389-0_23.

[5] C., S. Jensen; D., Lin, and B., C. Ooi, (2004). "Query and Update Efficient B+-Tree Based Indexing of Moving Objects", Proceedings of 30th International Conference on Very Large Data Bases, Toronto, 29th August-3rd September, pp. 768-779.

[6] M., F. Mokbel; X., Xiong, and W., G. Aref, (2004). "SINA: Scalable Incremental Processing of Continuous Queries in Spatio-Temporal Databases", Proceedings of 2004 ACM SIGMOD International Conference on Management of Data, Paris, 13th -18th June 2004.

[7] S., Prabhakar; Y., Xia; D., V. Kalashnikov; W., G. Aref, and S., E. Hambrusch, (2002). "Query Indexing and Velocity Constrained Indexing: Scalable Techniques for Continuous Queries on Moving Objects", IEEE Transactions on Computers, Volume: 51, Number: 10, pp. 1124-1140, doi:10.1109/TC.2002.1039840.

[8] Y., Tao; D., Papadias, and Q., Shen, (2002). "Continuous Nearest Neighbor Search", 28th International Conference on Very Large Data Bases, Hong Kong, 20th-23rd August 2002, doi:10.1016/B978-155860869-6/50033-0.

[9] X., Yu; K., Q. Pu, and N., Koudas, (2005). "Monitoring K-Nearest Neighbor Queries over Moving Objects", 21st International Conference on Data Engineering, Tokyo, 5th-8th April 2005, pp. 631-642.

[10] H., Hu; J., Xu, and D., L. Lee, (2005). "A Generic Framework for Monitoring Continuous Spatial Queries over Moving Objects", Proceedings of 2005 ACM SIGMOD Inter-national Conference on Management of Data, Baltimore, 13th-16th June 2005, doi:10.1145/1066157.1066212.

[11] D., A. Randell; A., G. Cohn, and Z., Cui, (1992). "Computing Transitivity Tables: A Challenge for Automated Theorem Provers", Lecture Notes in Computer Science, Volume: 607, pp. 786-790.

*****

A STUDY ON EMOTIONAL INTELLIGENCE AMONG INFORMATION TECHNOLOGY PROFESSIONALS IN COIMBATORE CITY, TAMIL NADU

Raies Hamid[116]

ABSTRACT

Emotion, in this most general definition, is a complex psychophysical process that arises spontaneously, rather than through conscious effort, and evokes either a positive or negative psychological response or physical expressions, often involuntary, related to feelings, perceptions or beliefs about elements, objects or relations between them, in the imagination. An emotion is often differentiated from a feeling. Emotional Intelligence (EI), often measured as an Emotional Intelligence Quotient (EQ), describes ability, capacity, or skill to perceive, assess, and manage the emotions of one's self, of others, and of groups. As a relatively new area of psychological research, the definition of EI is constantly changing. The study concentrates on the emotional intelligence level of IT professionals.

The study is important to IT professionals because emotional intelligence is being described as a new and better way of measuring an individual's chance of success in life. The software revolution totally changed the way we work. Probably no sector is untouched by information technology. As a result, IT industry employees not only the staff trained in computers but also professionals from all other fields. The rapid development of technologies crested totally new job categories. The emotional intelligence level of an IT professional will determine his ability to manage his/her own feelings and the feelings of others. It is said that higher the emotional intelligence, the greater the ability one's own feeling and to deal effectively with others. The aim of the study is to study the emotional intelligence among IT professionals.

KEYWORDS

Emotional Intelligence, Employees, Emotions, Emotional Intelligence Quotient, and IT Professionals etc.

INTRODUCTION

Emotion is complex, and the term has no single universally accepted definition. The study of emotions is part of psychology, neuroscience, and more recently, artificial intelligence. According to sloman, emotions are cognitive processes. Some authors emphasize the difference between human emotions and the affective behavior of animals. In Paul D. MacLean's classic triune brain model, emotions are defined as the responses of the Mammalian cortex. Emotion competes with even more instinctive responses from the Reptilian cortex and the more part of human decision-making and planning, and that the famous distinction made by Descartes between reason and emotion is not as clear as it seems. Emotions are thought to be related to activity in brain areas that direct our attention, motivate our behavior, and determine the significance of what is going on around us. Pioneering work by Broca (1878), Papez (1937), and MacLean (1952) suggested that emotion is related to a group of structures in the center of the brain called the limbic system, which includes the hypothalamus, cingulate cortex, hippocampi, and other structures. More recent research has shown that some of these limbic structures are not as directly related to emotion as others are, while some non-limbic structures have been found to be of greater emotional relevance. The following brain structures are currently thought to be most involved in emotion:

- Amygdala – the amygdala are too small, round structures located anterior to the hippocampi near the temporal poles. The amygdalae are involved in detecting and the learning what parts of our surroundings are important and have emotional significance. They are critical for the production of emotion, and may be particularly so for negative emotions, especially fear.

- Prefontal cortex – The term Prefontal cortex refers to the very form of the brain, behind the forehead and above the eyes. It appears to play a critical role in the regulation of emotion and behavior by anticipating consequences of our actions. The prefrontal cortex may play a important role in delayed gratification by maintaining emotions over time and organizing behavior toward specific goals.

- Anterior cingulated – The anterior cingulated cortex (ACC) is located in the middle of the brain, just behind the prefrontal cortex. The ACC is thought to play a central role in attention, and may be particularly important with regarded to conscious, subjective emotional awareness. This region of the brain may also play an important role in the initiation of motivated behavior.

- Ventral striatum – The ventral striatum is a group of sub cortical structures thought to play an important role in emotion and behavior. One part of the ventral striatum called the nucleus acumens is thought to be involved in the experience of goal-directed positive emotion. Individuals with addictions experience increased activity in this area when they encountered the object of their addiction.

- Insular – The insular cortex is thought to play a critical role in the bodily experience of emotion, as it is connected to the other brain structures that regulate the body's autonomic functions (Heart rate, breathing, digestion, etc).

Emotional Intelligence (EI), often measured as an **Emotional Intelligence Quotient (EQ),** describes ability, capacity, or skill to perceive, assess, and manage the emotions of one's self, of others, and of groups. As a relatively new area of psychological research, the definition of EI is constantly changing. The term emotional intelligence appears to have originated with Charles

[116] Research Scholar, Department of Management, Punjab Technical University, Punjab, hamidraies10@gmail.com

Darwin's origin of the species (1872), which speculated about the broader emotional social intelligence necessary for human survival and adaptation. As in the case in so many fields, language is a primary case for bias when it comes to determining a subject's origin or history. Payne many have been the first to coin the phrase for the Anglo-Saxon world, but, historically, he was preceded 20 years Carl Lanes, a Dutch science fiction writer of the early 1960s. in two novels, Lanes elaborates on the concepts underlying EI, even using the phrase emotional Quotient. These books were never translated, but they formed the base of an immensely popular radio show. India as emerged as the fastest growing IT hub in the world, its growth dominated by IT software and services such as Custom Application Development and Maintenance (CADM), System Integration, IT Consulting, Application Management, Infrastructure Management Services, Software Testing, Service-Oriented architecture and Web services.

OBJECTIVE OF STUDY

- To study the level of emotional intelligence among IT professionals.
- To study the relationship, if any, between age, gender, marital status, family size, annual Income, work experience, and emotional intelligence.
- To study the correlation between age family size, work experience, and emotional intelligence.
- To offer suggestions to the IT professionals to improve their emotional intelligence.

The samples selected for the study were IT professionals in and around Coimbatore City. Different professionals do require different levels of emotional intelligence. The selected professionals belong to age group of 22 to 38 years with work experience of 1 to 10 years. Initially the sample size decided was 250 employees but 120 were selected randomly using lot method. The data has been collected among the professionals of selected IT companies in the Coimbatore city. The researcher got the prior appointments from the professionals and in the specified time they were approached and the data were collected using the questionnaire. The chi-square test was used to find the relationship between the age, marital status, annual income, gender, work experience, family size to that of the emotional intelligence score of the respondents. Karl Pearson's Co-efficient of correlation was used to find the type of correlation between age, family size, work experience in years and the emotional intelligence score of the respondents.

ANALYSIS AND INTERPRETATION

It is found from the overall analysis that 60% of respondents belong to age group of 25 – 29 years. 73% of respondents belong to male gender. 72% of respondents are unmarried. 50% of respondents belong to family size of 3-4. 48% of respondents belong to annual income of 2-4 lakhs. 57% of respondents belong to the work experience of 2-4 years.

Table-1: Reliability of Scales

S. No.	Factors	No of items	Cronbaha Alpha
1	Emotional Intelligence	16	0.771

Sources: Primary Data.

Table-2: Reliability of Scales

S. No.	Emotional Intelligence	Number of Respondents	Percentage
1.	Low E.Q.	8	7
2.	Moderate E.Q.	72	60
3.	High E.Q.	34	28
4.	Extremely High E.Q.	6	5
	Total	120	100

Sources: Primary Data.

It is found from the overall scoring and analysis that 60% of respondents belong to moderate Emotional Intelligence Level. The first objective of the study is to find out the existence of relationship between age, gender, marital status, family size, annual Income, work experience, and emotional intelligence.

Emotional Intelligence (Chi – Square Table)

Table-2: Level of Significance Between Demographical Variables and Emotional Intelligence.

Factor	Calculated X^2 Value	Table Value	Degree of freedom	Remark
Age Group	7.16	16.919	9	Not significant at 5% level
Gender	18.37	16.919	9	Significant at 5% level
Marital Status	7.66	7.815	3	Not significant at 5%level
Family Size	19.72	16.919	9	Significant at 5% level
Annual Income	9.22	16.919	9	Not significant at 5% level
Work Experience	5.53	16.919	9	Not Signification at 5% level

Sources: Primary Data.

Table-3: Correlation Between Between Age and Emotional Intelligence of Respondents.

Factor	Spearman Correlation	Lamda	Asymp. Sig.	Level of correlation
Age	0.1336	.16	0.000*	Low correlation
Family Size	0.1654	.28	0.000*	Low correlation
Work Experience	0.1231	.14	0.000*	Low correlation

Sources: Primary Data.
*Significant at 5% Level

From the above table it can be inferred that the R value is found to be low for demographical variables (Age, Family size, Work experience) hence is it clear that the level of correlation between emotional intelligence and demographical variables are too be low.

RESULTS AND DISCUSSIONS

Self- awareness (knowing and understanding your own feelings)

Focus on a particular area you're having difficulty with – anxiety or anger, for example. At the beginning or end of the day write down what triggered them? What made you feel better? Keep a journal and you"ll start seeing a pattern – now you 're ready for self regulation.

Self-regulation (Being able to control your emotions)

You must take responsibility for your own feelings and stop blaming others. Replace "they've made me feel this way" with "what can I do to change the way I feel?" Blaming others only gives your power away!

Motivation (persistence when things get though)

Make sure you are moving towards a goal rather away from something. When we force ourselves to do something "we have to or else". But if we're really motivated, we think "I want to because …." So work out want you really want and why.

Empathy (The ability to sense and respond to others' emotions)

Try imaging what it's like to be that other person. What would you do, say or feel if you had their perspective on the situation? This is key to influencing.

Social Skills (Influencing and handling other people's emotions)

Before you communicate with someone, work out what you are trying to achieve. If you've thought about it, you're more likely to achieve it.

SUGGESTIONS

- Take responsibility for your emotions and your happiness.
- Examine your own feelings rather than the actions or motives of other people.
- Develop constructive coping skills for specific moods. Learn to relax when your emotions are running high and to get up and move when you are feeling down.
- Make hunting for the silver lining a game. Look for the humor of life lesson in a negative situation.
- Be honest with yourself. Acknowledge your negative feelings, look for their source, and come up with a way to solve the underlying problem.
- Show respect by respecting other people's fee lings.
- Avoid people who invalidate you or don't respect your feelings.
- Listen twice as much as you speak.
- Pay attention to non-verbal communication. We communicate with our whole selves. Watch faces, listen to tone of voice, and take note of body language.

Realize that improving your emotional intelligence will take time and patience.

CONCLUSIONS

Emotional intelligence is the capacity to create positive outcomes in your relationships with others and with yourself. Positive outcome include joy, optimism and success in work and life. Thus emotional intelligence involves the whole gamut of internal and external emotional management leading to the fruitful existence of an individual in this universe increasing emotional intelligence, feelings, marriage and health.

Thus emotional intelligence is very essential because it is defined as skill to know whom you are. How your thoughts? Feelings and actions inter – related.

In this study emotional intelligence of I.T. professionals are analyzed and it is found that majority of respondents belong to moderate E.Q. level. There is no significant relationship between age, marital status, annual income, work experience and emotional intelligence of the respondents.

There is a significant relationship between gender, family size and emotional intelligence of the respondents. There is low correlation between age, family size, work experience and emotional intelligence of the respondent. Suggestions are offered by the researcher to improve the level of emotional intelligence among the I.T. professionals.

REFERENCES

[1] Mc., Shane, (2050. Organizational Behaviors, New Delhi; Tata McGraw Hill.

[2] Stephen., P. Robbins, (1995). Organizational Behaviors Concepts, Controversies and Applications, New Delhi: Prentice Hall India Private Limited.

[3] James, Hassett, and Kathleen, M. White, Psychology in Perspective, New York; Harper and Row Publishers

[4] Salovey, Peter, and Sluyter, David, (1997). Emotional Development and Emotional Intelligence, Educational Implications, Basic Book.

[5] Singh, Dalip, (2006). Emotional Intelligence at Work, New Delhi: Sage Publications India Pvt. Ltd.

[6] Kothari, C. R., (2006). Research Methology, Methods and Techniques, New Delhi: New Age International (P) Limited, New Delhi.

[7] Gupta, S. P., (1969). Statistical Methods, New Delhi: Suthan Chand & Sons..

*****

GUERILLA MARKETING AND ITS WEAPONS IN ADVERTISING ERA

E. Krithika[117] S. Parimala[118]

ABSTRACT

Guerrilla marketing is an advertising strategy, in which low-cost unconventional means are utilized, often in a localized fashion or large network of individual cells, to convey or promote a product or an idea. The term guerrilla marketing is easily traced to guerrilla warfare which utilizes a typical tactics to achieve a goal in a competitive and unforgiving environment. The unconventional marketing involved tries to get the most out something small, and make a lasting brand image in the consumer's mind. The objective of guerrilla marketing is to create a unique, engaging and thought-provoking concept to generate buzz, and consequently turn viral. The term was coined and defined by Jay Conrad Levinson in his book Guerrilla Marketing. The term has since entered the popular vocabulary and marketing textbooks. Guerrilla marketing involves unusual approaches such as intercept encounters in public places, street giveaways of products, PR stunts, or any unconventional marketing intended to get maximum results from minimal resources. More innovative approaches to Guerrilla marketing now utilize mobile digital technologies to engage the consumer and create a memorable brand experience. Guerrilla marketing focuses on low cost creative strategies of marketing. Basic requirements are time, energy, and imagination and not money. Sales do not compose of the primary static to measure business but is replaced by profit. Emphasis is on retaining existing customers then acquiring new ones. And the paper focuses on marketing through Guerrilla and its techniques.

KEYWORDS

Advertisement, Creativity, Viral, Ambush, Guerilla, Strategies, Marketing, and Media etc.

INTRODUCTION

Guerilla Marketing is a brilliant idea, involving the customer in a surprising, unconventional marketing activity. The term "Guerilla" (battle) roots back to the war of independence in Spain and Portugal, the revolution in Cuba, and the Vietnamese War. Here "Guerilla" stood for an attack strategy based on the surprise effect and on acts of sabotage, which was used by smaller groups that stood against a massive military force.Guerilla Marketing – an alternative marketing form that has been experiencing increasing importance in the advertising landscape, ever since customers are besieged with classical marketing communication via the traditional channels TV, Magazines, Radio and Direct Mail.

GUERILLA ENTERS INTO MARKETING

In the **1960's** US firms brought the "Guerilla" tactics to Marketing, when they needed new ways to exceed competitors. Then they merely attacked weak points of competitors by implementing preliminary injunctions for their campaigns for instance. Only in **1983** did Jay Conrad Levinson make the term "Guerilla Marketing" known as a philosophy for small and start-up companies to successfully market their business with a small amount of money. He based the success of a marketing strategy on the use of non-traditional marketing channels, customer proximity, insistency, and patience.

DEVELOPMENT OF GUERILLA MARKETING

Technology Development

"Markets today are changing fast. Price-sensitive customers, new competitors, new distribution channels, new communication channels, the Internet, wireless commerce, globalization, deregulation, privatization… the list goes on. And it is not only markets that are changing, but technologies that support them: e-commerce, e-mail, mobile phones, fax machines, sales, and marketing automation, cable TV, videoconferencing. It is imperative that companies think through the revolutionary impact of these new technologies." Through the drastic technological development until today and its complimentary change in the advertising market, Guerilla Marketing has developed into a marketing form mostly used for Promotion these days. New forms such as **Ambush** and **Viral Marketing** have evolved. The evolved Guerilla Marketing form is now used by companies of all sizes; Global Players such as BMW use it, and start-up companies do so to create brand awareness. How such Guerilla Marketing activities are perceived by the actual prospect was researched through a customer survey. In fact a majority of the respondents stated that such advertisements do catch their attention and interest, whereas they consciously avoid traditional marketing as for instance TV advertisement.

Guerilla Marketing can find a way to actually reach the customer, at least in the first two steps of the AIDA Model. This great potential of the Guerilla Marketing was confirmed by Marketing Experts in the course of an interview. They could imagine that the traditional Marketing Channels such as TV and magazine, who will also be at the forefront in the future, will take over characteristics of Guerilla Marketing in order to actually reach the customer. Financially strong companies should use Guerilla Marketing as a complimentary tool to present the brand in a multimodal way. Smaller firms can make use of the cost-effective strategy to get the spotlight on their brand. Good co-operation with for instance the Public Relations department is essential to increase the Guerilla Marketing effect enormously. However its potential is used, it has to be done cleverly to stand out from the 3000 advertising messages a day with which today's consumer is flooded. Guerilla Marketing will adapt with the customers, find ways to surprise them and interest them by putting the idea in the forefront, not the brand. This development leads to some interesting questions:

- Is this trend going to continue?
- Will all companies be able to use it in the future?
- Will customers still be surprised and show the desirable reactions even though many companies already use Guerilla Marketing today?
- Do small companies have a chance with their Guerilla Marketing approach when larger Companies

[117] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, krithikaelango.18@gmail.com
[118] Research Supervisor, Ana University, Tamil Nadu, India, dr_parimela@gmail.com

are willing to spend a fortune on their Guerilla Marketing activities?

GUERILLA MARKETING IN MARKETING MIX

Guerilla Marketing changed over the years. In the past it was a business philosophy that influenced all aspects of the business equally. Today Guerilla Marketing is often only used in the form of a campaign. Companies often operate according to their traditional marketing philosophy, but Guerilla Marketing for single campaigns is different. There the balance in the marketing mix shifts towards one of the 4 P's. Figure2 shows that 70% of the campaigns put their focus on promotion. The remaining 30% place their focus equally on price, place, and product.

Figure-1: Guerilla Marketing for Single Campaign

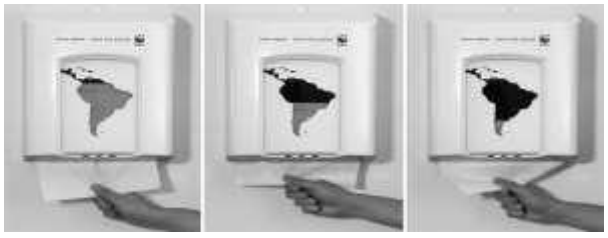
Sources: Authors Compilation.

The following examples show campaigns where the focus is on one of the 4 P's. Here it is essential to outline that sometimes only one of these 10% of the Marketing Mix is designed according to the Guerilla Marketing principles. The remaining 90% can belong to a traditional Marketing approach.

Guerilla Promotion (70%)

In most cases, Guerilla Marketing appears in the form of promotion. Even though Guerilla Marketing tries to be different than the traditional marketing strategies, it uses the same channels to bring the message across such as public relations, advertising, sales promotion, or direct mail as well as the specific instruments described before. Actually, the possibilities do not have limits.

Figure-2: Save Paper- Save the Planet



Sources: Worldwildlife.

The non-profit organization WMF supplies one of the numerous examples of Guerilla Marketing promotion with its slogan: **"save paper – save the planet".** By pulling the paper towel out of the box, the level of the overall paper towel pile decreases. The user cannot only see the decreasing pile, but also the effects that the use of the paper towels has on the rain forests in South America. With this simple example WMF tries to redirect the attention of the user towards WMF's interest to maintain the rain forests.

GUERILLA PRICING (10%)

The focus of Guerilla Marketing strategies can also be pricing. Even though we live in a world where prices play a big role within the purchase decision, only 10% of all Guerilla activities focus on price. It describes a new way to differentiate itself from the rest of the competition. A competitive offer is still a success factor for some campaigns as the following witty idea shows. Customers of the electronic retailer Media Market were promised to get back their money if they bought TV's before the Soccer European Championship in case that the German soccer team won the tournament. Many people took the opportunity and bought one or more TV's. In the end Germany did not even get into the final round and the customers did not receive their money back. Media Markt was the winner.

GUERILLA DISTRIBUTING (PLACE 10%)

The distribution of a product can also be a special experience for the customer. A good distribution can boost sales tremendously. Even though books have a hard time to survive in competition with to other media, the famous Harry Potter series by J. K. Rowling showed that even books can make the day of retailers and delivery services. Especially young readers were waiting for the new book release and wanted to buy the book at the day of publication. Book shops, retailers, and delivery services in many countries offered customers to buy or to receive the book at midnight in order to be one of the first to own the book. A lot of fans even came in costumes to book stores to buy the first copies. Others were willing to pay extra to get the book delivered to their homes at midnight.

GUERILLA PRODUCTING (10%)

Also the packaging, the form, and the brand can be a vehicle for Guerilla Marketing. One example introduced Kellogg's with its Drink 'n Crunch cup shown in the picture below:

Figure-3: Kellogg's Drink

Sources: Business Week.

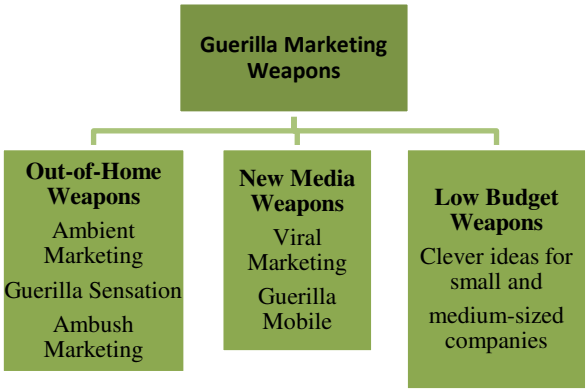
The cup allows the customer to have cornflakes-to-go. The one-person-portion in the cup only needs milk in addition to make it a snack that can be enjoyed not only at home. This makes the cornflakes an alternative to chocolate bars and other sweets. Kellogg designed this portable cereal cup. Pop the top and add the milk, and it never mixes with the cereal until they meet in your mouth. The cup is about $2

GUERILLA MARKETING INSTRUMENTS

Over the years Guerilla Marketing has developed and created many forms, so called weapons. These weapons should support

the root philosophy of Guerilla Marketing which entails creative, witty inspiration. As we know Guerilla Marketing is a dynamic way to create brand awareness. Done with simple tools those actions should be spectacular and clever to catch the attention of a certain target group. To get an overall picture of the potential of Guerilla Marketing weapons we have classified them by categories: Out-of-Home, New Media, and Low Budget Weapons. Again no official definition exists, but the following distinction by experts of Guerilla Marketing Portal (GMP) has been build up logically and is therefore referred to in this paper.

Figure-4: Guerilla Weapons



Sources: Authors Compilation.

OUT-OF-HOME WEAPONS

As the name "out-of-home" suggests, these weapons refer to marketing activities that are actually realized at public locations. At best it does not only catch the interest of people who pass by, but media interest as well. Newspaper reports about the action can create extra publicity for the advertised company and stimulates that people talk about the product. The expansion of brand awareness is the aim. The most successful weapons in the category out-of-home are Guerilla Sensation, Ambient Media, and Ambush Marketing.

Ambient Marketing:

The term became well-known in the 1990s. Ambient Media refers to non-traditional out-of-home advertising. While other out-of-home Marketers advertise on large-scale billboards, ambient advertisements are posted on manhole covers, cranes, pizza cartons, free postcards in bars and so on. They are all a little more unusual displays. Until today newer ambient media have come up, such as messages on large-format screens in subways or handles of supermarket trolleys. There are no limits to formats there.15 The American Mark Hughes had what is probably one of the most profitable ideas concerning the placement of an advertisement. He used the blank backside of the fortune cookie slips given out in Chinese restaurants for an advertising message. These – in production cheap - marketing messages reached 7 million restaurant guests in one week and were therefore very profitable. Mark Hughes has been able to make a lot of money by selling this noticed advertising space to companies. The important aspect is to seek out the target group at their preferred location in an entertaining way. Ambient Marketing tries to approach the customer individually through the place of contact. Some experts claim that Ambient Marketing does not belong to Guerilla Marketing, since it does not fulfill the typical characteristics of being surprising and dynamic. Indeed Ambient is more static, one-sided marketing. Still it can also be a medium with good potential to bring attention to a brand or product in a creative way.

Figure-5: Big Bazaar Guerrilla Marketing

Sources: India Reviews.

Indian Example: Big Bazaar, Pantaloons, Future Bazaar, eZone are all part of this group and they are taking on the biggies like Shoppers Stop, Lifestyle, and Tata's Westside. In order to do the same, Future Group has come up with 3 catchy/cocky and cheeky ad campaigns which surely do catch your eyes.

Westsside: Make a smart choice!
Shoppers! Stop. Make a smart choice!
Change Your Lifestyle. Make a smart choice!

Guerilla Sensation:

Guerilla Sensation is very similar to Ambient Marketing. Therefore it is easier to show the difference. As described above, Ambient Marketing positions advertising at unusual places. Hereby the main focus is not necessarily on the idea, but on the advertising space itself. People are confronted with advertisements where and when they do not expect it. In general Guerilla Sensation works with the same principle, but it is only used on a very limited number of events and activities. The number of prospects that are exposed to the advertisement is therefore relatively small compared to a few million people that might get in contact with an Ambient Marketing campaign as the example of the Fortune Cookies described above. But such Guerilla Sensation actions can gain further, non-regional attention by making it to the media. Often newspapers are looking for extraordinary, unusual pictures, which are often delivered by Guerilla Sensation activities. Through good Public Relations controversial or very original Guerilla ideas of a brand can be a widely discussed issue in the media as well. Of course the line between Ambient Marketing and Guerilla Sensation is thin. Some Guerilla Sensation activities become an Ambient Marketing activity, because modern technologies spread the message. Guerilla Sensation characteristics are those described in the Guerilla Marketing definition.

Ambush Marketing:

Ambush means attack out of the blue, reminding at the Guerilla Attacks from Che Guevara. It stands for a sneaky out-of-home

marketing method, which promotes a brand at huge events without paying a sponsorship fee. At many major events one brand of a particular category pays a high price to be the exclusive sponsor, which leaves their competitors be left in the dark. Ambush Marketers then still find a way to make notice of their brand in connection with the event, since it attracts the attention of thousands of visitors and even viewers on TV. Pepsi for instance placed a huge oversize Pepsi bottle close to a soccer game which was sponsored by Coca Cola. Giving out company material on a fair without having a stand is also Ambush Marketing. The legal boundaries can be thin as Vodafone experienced in Australia in 2002. The rugby match between the All Blacks from New Zealand and the Wallabies from Australia was interrupted by two naked streakers whose bodies were painted with Vodafone logos. The bizarre details unsheathed the wrongful behavior of Vodafone. The CEO of the Australian Vodafone division knew that something provocative would happen and agreed to the proposal of an anonymous caller. He even agreed to take care of all arising costs and legal issues. Furthermore Vodafone was the sponsor of the participating Australian team, but his main competitor Telstra gave the arena its name. Apart from the legal difficulties that Vodafone had to face, the success of the campaign was arguable, since the match might have been influenced by the interruption and upset many fans.

Figure-6: Reliance's Ambush Marketing

Sources: India Reviews.

Indian Example: To promote RGI's health insurance products, people wearing armour roamed crowded areas, with the message "or call us" displayed on their backs. The point of the message: To protect yourself from health hazards such a diseases or accidents, go in for insurance, not armour.

NEW MEDIA WEAPONS

New technologies change our lives and they often make it easier due to mobile phones, internet, unlimited information, and shopping possibilities that enable customers to access the resources of the world with a click on the computer mouse. Of course this also gives businesses the possibility to use the advantages that technology provides. Two very strong instruments that use the modern possibilities are described below: Viral Marketing and Guerilla Mobile.

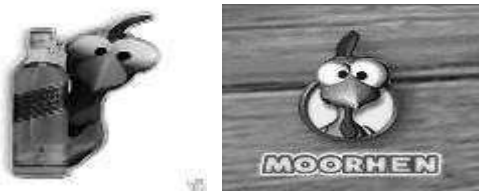
Guerilla Mobile:

Not only has the PC offered unlimited possibilities to marketers. Since the number of mobile phones exceeds the number of inhabitants in many countries, the cell phone is a permanent companion of prospects. Therefore it was only a matter of time until marketers took the opportunity to reach customers and prospects at any place at any time. The wireless connection provides the possibility to present marketing messages in different ways via SMS, MMS, Bluetooth, or Infrared.

Viral Marketing:

Viral Marketing is a very important weapon of Guerilla Marketing. The beginning of Viral Marketing in Germany is often linked to the computer game "Johnnie Walker Moorhuhn", where the player gets to shoot grouses (German: "Moorhuhn") in the Scottish highlands. The game was developed to promote the Scotch Whiskey brand Johnnie Walker in the German gastronomy sector. In 1999 it was shown in some German bars and then offered as a free download on the internet. Unexpectedly the game became so popular that people of all ages actively played it. The clue which makes this example Viral Marketing is that the game players recommended "Moorhuhn" to friends by forwarding the link or by word-of-mouth. The game downloads had exponentially increased within a very short time and the brand name Johnnie Walker was spread free of charge to millions of individuals.

Figure-7: Johnnie Walker Moorhuhn Game – Facebook



Sources: Moorhens.

The above example vividly shows how successful Viral Marketing can be. The idea is to get an advertising message to as many contacts as possible and as fast as possible. This is realized by human multipliers who are animated to pass the message along for free. The challenge for a Viral Marketer is to build a motivation in a message for people to spread it. If that is successful, the snowball effect starts and the potential for the viral message's exposure and influence is produced. The message rapidly spreads, similar to the way email jokes and software viruses are sent around to the online community. It has to be mentioned that Viral Marketing is not a computer virus, it is harmless. Encouraging people to pass the message along to others for free is an art. A personalized message or a motivational prize is often successful tools. In any way, the campaigns should be smart and extraordinary, since a conventional advertisement might not be accepted and spread. Winning is often the usage of sex, humour, scandal, and provocation. The Viral Marketing concept is not dependent on a certain media. Communication between two persons has always existed in various ways. Then the message often spread by word-of-mouth, but through the internet Viral Marketing has experienced a boom. The impact, coverage, and pace that this media offers today is astonishing. Especially, Hotmail is often mentioned as the showcase for Viral Marketing. This free email service by Microsoft was one of the first to gain remarkably from Viral Marketing. Hotmail's brand awareness and market share increased rapidly, when they started to add a frank message to the footer of all E-Mails sent out by Hotmail users.

For **example:** "*Join the world's largest free e-mail service with MSN Hotmail.* http://www.hotmail.com. 20 Hotmail users thus automatically became sales promoters for their e-mail service provider. From its start in 1996 to its 12 millionth user, hotmail only spend 500.000 USD on advertising, while competitors spent 20 Million US-Dollar and did not nearly manage to attract that many customers. Other successful viral formats are for instance Facebook applications and viral videos on YouTube.

Figure-8: Creating Promotional Pages through Facebook

Sources: Facebook.

The Pears product was exhibited by creating a page on facebook, and a 3D card was created on their page. The card contains "Like to unlock" where people will like the product automatically. And this spreads widely and cheaper at cost.

LOW BUDGET WEAPONS

This Weapon refers to Guerilla Marketing for new, small, and medium-sized companies, who only possess a small marketing budget, like Levinson already pointed out in the 1980's that does not necessarily mean that those companies have disadvantage compared to financially strong competitors. But since their capital is low, the top priority is to use it as efficiently as possible. Guerilla Marketing should put this into practice by focusing on local culture with its geographical, sportive, social network, its ritual, needs, habits, norms, traditions, and values. Clever ideas appear through unconventional methods which are supposed to catch attention of the target group.

GM STRATAGEM OR LINE OF ATTACH

Creativity

In guerilla marketing, the world is your billboard / 30-second TV spot / lower-third. Guerilla marketing begins with out-of-the-box ideation. To develop on-target, unexpected, effective guerilla campaigns, throw convention to the wind and get your creative juices boiling. The heart of every successful guerilla campaign is creativity. Without it, you're conventional or worse yet, annoying

Unexpected

A billboard is expected. On your way to work, there are always the same billboards—with bright colors, maybe a catchy line that makes you smile. It's expected.

Figure-9: German Houseware Manufactures - Miele

Sources: Sparxoo Branding Experts for Digital Era.

The German housewares manufacturer, Miele created a campaign that transformed a typically static flat billboard surface into so much more. To satirically demonstrate the power of their vacuum's suction, they draped an actual hot air balloon over a billboard with their vacuum cleaner sucking the balloon into the billboard.

More with Less

When brainstorming guerilla initiatives, it's easy to become fantastical and impractical. If you're a smaller company with budgetary constraints, pasting a 20-foot soccer ball to the side of building might not be best consideration. Being reasonable, intelligent and unexpected stretches your creativity to develop a very effective guerilla campaign without an exorbitant budget. You don't need to be Superman to have a Superman idea.

Figure-10: A Yoga Center Ad

Sources: Ads of the world

To advertise the Superman brand, small posters were pasted to lampposts to give the impression they were tied in a knot. Or, something as simple as a straw can be valuable ad real estate for a yoga company.

Maximize Your Surroundings

Lining the street with coffee cups does no one good. It's not intelligent, too abstract, not relatable and a pedestrian inconvenience, but if you utilize your environment in an unexpected way.

Figure-11: Folgers Innovative Way

Sources: Poshima – Smart Web Publishing System.

Intelligent way–like Folgers, when they used the steam from manhole covers to show the steam from their coffee–you're making more meaningful impressions.

Interactive

Though interactivity is not a staple in all guerilla campaigns, it brings consumer / company relationship to more meaningful space.

Figure-12: Ikea Luxury Appeal- Swedish Company

Sources: Freshome.

Take example from Swedish innovator and furniture phenomenon, Ikea. To demonstrate the appeal of an "Ikea living room," they transformed a bus stop into a place you wouldn't mind relaxing to enjoy a couple Swedish meatballs. They beautified a bus stop with their furniture - making it not only an interactive display, but a cozy interactive display.

GUERILLA MARKETING TODAY

The main reason why companies opt for Guerilla Marketing these days can be differentiated between small and larger organisations. Smaller companies obviously chose the cost effective principle since they only have a small advertising budget on-hand. Larger, more solvent companies are often able to afford TV and Print Ads. They use Guerilla Marketing since it achieves value that cannot be created by classical advertising alone.

OBSTACLES OF GUERILLA MARKETING

Guerilla Marketing goes another way than traditional marketing. Therefore it is often difficult to recognize the fine line between provocation and offence, between drawing the attention to a company and to create a negative reputation, to get new customers and to lose prospects. The following example shows a Guerilla Marketing campaign that failed:

There are sometimes invisible limits that a company should not cross: A company promoted its new comedy series by placing a large electronic billboard featuring an adult character in the city centre of Boston, MA. Lose cables were hanging out of the billboard. People passing by believed that they were explosives and panicked. They called the police and an anti-terror unit was sent to disarm the billboard which was wrongly interpreted as a bomb. The city centre and all access roads were closed. Later that day the mistake was uncovered and it resulted in two arrests and costs of 500,000 US$. Even though the advertising was already in the city for some days and was also placed in other major US cities and did not cause any excitement there, it is obvious that those in charge should have dealt with this topic in a more sensitive way.

In times of Viral Marketing and media coverage from all over the world, it is almost impossible to let something "disappear" that was presented to the public once before. Consequently this means that every publication that has not been thought through carefully can damage the reputation of company and can result in the loss of customers.

CONCLUSIONS

In the future consumers will take a negative view of advertising messages, if those are inartificial, boring, exchangeable, and without strong ideas. Clever, funny, or original messages are needed, since ideas will always have an economic cycle. Against the general perception, Guerilla Marketing is not a trend but an instrument. It has existed for over 50 years now and gained importance since its invention. Especially in the B2B sector Guerilla Marketing is applied more and more. In the future strategists should think and act modern. They should not disregard Guerilla Marketing, but realise that classical advertising does not reach many target groups. Forms of approach need to be found that are accepted by the audience and not ignored. Companies have to be sensitive and pay close attention to their target group.

REFERENCES

[1] Jay, Conrad, and Levison, David ePerry, Guerrilla Marketing, 4th edition: "Easy and Inexpensive Strategies for Making Big Profits from Your Small Business**".**

.

[2] E., Krithika, (2011). "Guerilla Marketing", In: Tamil Nadu Engineering College, 21st December 2011, Coimbatore, TamilNadu, Unpublished Paper.

[3] Ikea, (2010). Examples: "How Ikea Swedish Furnishing Transformed The Bus Stops", viewed on December 29th, 2011, http://freshome.com/ikea-parisian-saga-continues-with-12-bus-shelters-turned-into-livingrooms/

[4] Strategies, (2011). "Guerilla Marketing Strategies", Viewed on November 05th, 2011, http://sparxoo.com/top-5-guerilla-marketing-strategies/

[5] Advertisements (2010): Examples "Kellogg's Nuty Crunch Package Design Selected as the Best Package Design", Viewed on December 15th, 2011, http://indiareviews.wordpress.com

*****

ENTREPRENURIAL COMPETENCY COMPONENTS OF SHG WOMEN ENTREPRENEURS – FACTOR ANALYSIS

M. R. Vanithmani[119]

ABSTRACT

The spirit of entrepreneurship helps individuals to practice the art of innovating ideas, products and services out of their knowledge, skills, competencies and other resources available to them. Further entrepreneurship enables them to optimize the use of individual's resources to efficiently organize and manage innovations. Thus, entrepreneurship is primarily human resource development concept, starting with sensing opportunity, generating ideas, products and processes, assessing and mobilizing resources & finally organizing and managing innovations to solve problems.

Through entrepreneurship training/education, growth is stimulated to make human resources productive for personal as well as societal benefit. Due to lack of capacity building, development process remains incomplete among women, which can be completed through entrepreneurship training/ education, motivating them for excellence, understanding behavioral competencies, developing and strengthening entrepreneurial qualities, heightening innovation and creativity, planning and resourcing support for success, growth influence, and leadership and organizing competencies.

This study based on these factors aimed at deriving the factors which contributes to entrepreneurial competencies of women entrepreneurs. Training programmes aimed at the Self Help Group women shall be planned in such a way to knob the entrepreneurial competencies.

KEYWORDS

Entrepreneurial Competency, Self Help Groups, Trainings, and Women Entrepreneurs etc.

INTRODUCTION

Throughout history, people have formed groups with others who have something in common with them, and oppressed people have joined together to overcome the conditions they face. The self-help group is a small body formed by the people for meeting their specific objectives, particularly credit. Self-help groups build on the strengths of their members. Self-help groups (SHGs) play today a major role in poverty alleviation in rural India. A growing number of poor people (mostly women) in various parts of India are members of SHGs and actively engage in savings (in actual term Thrift) and credit, as well as in other activities (income generation, natural resources management, literacy, child care and nutrition, etc.).

Many NGOs are promoting the SHG mechanism and linking it to various other development interventions. Whereas there is ample evidence that the SHG approach is a very effective, efficient and relevant tool for organizing and empowering the poor, do arise with design, development and introduction of programmes to promote income-generating activities (IGAs) that will generate sufficient, sustainable and regular income. In contrast to formal organization, self-help is highly personal, no-hierarchical and without division of labour. Self-help favors experience over expertise.

A typical rural women's self-help group is a good example of capacity building for prospective entrepreneurs. It aims at enabling members with no educational or industrial or entrepreneurial background to become self-dependent and self-reliant by developing and enhancing the decision-making capacity of members and instilling in them the strength and confidence for solving their problems.

In the present study, the researcher has focused on the analysis and interpretation of entrepreneurial competency components of women entrepreneurs in SHGs, which triggers them toward entrepreneurial success. Based on theoretical framework linking Entrepreneurial Competency and factors affecting it, the study was designed to investigate the components of entrepreneurial competency of women entrepreneurs in Self Help Groups.

LITERATURE REVIEW

Competency is concerned with the long-term performance of a firm, an industry or a country related to its competitors (Ramasamy, 1995; Feurer& Chahar Baghi, 1994). It is also a multi- dimensional concept, including not only performance, but also potential and the process of generating performance (Buckley et al, 1998; World Competitiveness Report, 1993).

Azad (1989) in his research on "Development of entrepreneurship among Indian women" stated that there was lack of adequate theoretical understanding of the psychological characteristics, creates a severe gap, that needs to be bridged through proper Entrepreneurial Development Programmes. The author pointed out the following aspects as motivational factors which makes women to plunge in entrepreneurial career as, economic compulsion, use of knowledge and skills, need for achievement , success of others and frustration in previous job or occupation.

Prince & Fleming (1991)in their study on minority and women business owners in Colorado, USA, found that business owners who completed the fast track training programmes have been positively affected and would be much more successful in the future because of it. They ascertained that, graduates have developed growth plans for their businesses and the respondents had a hedge against failure and consecutively their businesses were converted form high –risk ventures to sound ventures.

Further the qualitative analysis done by Tomas.W.Y. Man & Theresa lace (2000) about "Entrepreneurial competency of SME owners / Managers in the Hong Kong Service Sector", revealed the relationship between behavioural bases and competency. The competency approach is a way of studying individual characteristics leading to the accomplishment of job role of organisational success. It has been widely applied to the study of managerial performance since the work of Boyatzis

[119] Director (MBA & HRD Cell), Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, vanithaselvakumar@gmail.com

(1982) and increasingly in field of entrepreneurial performance. By using various qualitative techniques, many studies have been conducted to identify different entrepreneurial competency in different contexts. (Adam & Chell, 1993; Bird, 1995; Darkan et al ,1993; Hunt et al,1998; Lau et al,1999; Mc.Clelland, 1987 ; Snell & Lau 1994 ; Murray, 1996 ; Marain & Staines,1994). Thompson *et al*'s (1996) have studied the competence of SME top team members by consideration of theoretical development in organizational success under the influence of organizational culture.

RESEARCH METHODOLOGY

The research design is descriptive in nature as the study is aimed at undertaking in-depth analysis on various parameters related to entrepreneurial competency of self help group women entrepreneurs.

25 SHGs representing each taluk of Coimbatore district of Tamilnadu state in India i.e., Coimbatore (North & South), Mettupalayam, Sulur and Pollachi were selected. From each group 4 members were selected i.e., 1 leader/ facilitator and 3 members, making a sample of 100 per taluk, constituting an overall sample size of 400.

Respondents were selected randomly from Self Help Groups from 4 taluks using **Snowball Sampling Method,** in which one respondent give lead to others (Thomas W.Y.Man and Theresa Lau, 2000). Initial contacts got through NGOs were used to reach all other Self Help Group members at different geographical locations based on their personal links and networks.

ANALYSIS AND INTERPRETATIONS

The spirit of entrepreneurship helps individuals to practice the art of innovating ideas, products and services out of their knowledge, skills, competencies and other resources available to them.

Further entrepreneurship enables them to optimize the use of individual's resources to efficiently organize and manage innovations.

For SHG women entrepreneurs' entrepreneurial opportunities are flooding through supporting institutions and governmental schemes. Right from the socio economic factors all other external and internal factors affect entrepreneurial performance of SHG women entrepreneurs.

Table-1: Demographic Distribution of Respondents

Age Wise Distribution			Based on Educational Qualification			Based on Marital Status		
Age Group	**Nos.**	**%**	**Qualification**	**Nos.**	**%**	**Marital Status**	**Nos.**	**%**
< 30 years	86	22	Up to School Level	267	67	Married	365	91
30-40 years	195	49	Technical or Diploma level	54	14	Un married	35	9
41-50 years	82	21	Up to college level	24	6	-	-	-
>50 years	37	9	No formal Education	54	13	-	-	-
Total	400	100	Total	400	100	Total	400	100

Sources: Primary Data.

Based on the analysis of data collected and interpretations made out of it, the demographic details of the respondents are given in **Table 1**. It is observed from Table 1, that majority of the respondents are in the age group of 31-40 year. This age group represents the economically productive women folk in the study area. And majority of the respondents have completed their basic school education. A minimum level of education is absolute essential for the active participation of women in the development programmes. In fact, education is one of the most essential inputs of the rural development. Majority of the respondents have at least school level education. Out of the 400 sample respondents surveyed, 91.25% of the respondents are married.

Table-2: Previous Entrepreneurial / Occupational Experience

Sl. No.	Previous Job	Number of Respondents	Percentage	Occupations	Number of Respondents	Percentage
1.	Own Business	165	41.30	Bakery	49	12.24
2.	Bakery	61	15.22	Business	144	35.92
3.	Coolie	49	12.17	Artisans	70	17.55
4.	Tailors	56	13.91	Employed	57	14.29
5.	Computer Centre	38	9.57	Weavers	20	4.90
6.	Others	31	7.83	Vegetable Vendors	8	2.04
				Others	52	13.06
	Total	**400**	**100.00**	**Total**	**400**	**100.00**

Sources: Primary Data.

Table2 depicts that among the 400 sample respondents, most of the i.e., 41.30% of the respondents have opined that they have previous entrepreneurial experiences before joining the SHGs. Rest of them have work experience in some skilled areas like in bakery, tailoring, computer operation, and general labours.

The remaining 7.83% of the respondents have gained experiences in areas like soft toy making, catering etc., that are categorized as others. Thus, it is found that most of the i.e., 41.30% of respondents have previous entrepreneurial experience before joining in SHG.

Among the sample population surveyed, most of the i.e., 35.92% of respondents family profession is business. Followed by it, 17.55% of the respondents' family members are Artisans. 14.29% of their family members are employed. 13.06% of their family is doing other occupations like seasonal businesses. 12.24% of respondent's family members are workers in bakeries.

Moreover, 4.905 of the respondents' families are weavers. Rests of 2.04% of their family members are vegetable vendors. Thus, it is found that primarily i.e., 35.92% of respondent's family members are into business.

Empirical evidence shows that women contribute significantly to family businesses mostly in form of unpaid effort and skills. The value of their efforts is underestimated both by families that take it for granted and by the societies.

Table-3: Nature of Business Performed by SHGs

Sl. No	Nature of Business	Number of Respondents	Percentage
1.	Manufacture	79	19.75
2.	Service	116	28.75
3.	Retail	129	32.00
4.	Wholesale	46	11.50
5.	Others	32	8.00
	Total	400	100.00

Sources: Primary Data.

The above table3 presents that out of 400 sample population surveyed; 32% of the SHG members are doing retail business. Followed by it, 28.75 per cent of them are in service activities. 19.75% of the respondents have expressed that their SHG is involved in manufacturing business. 11.50% of respondents are in wholesale business. The remaining 8 per cent of the respondents have opined that their SHGs are involved in other allied sector activities like fishery, dairy, piggery, poultry, cow, ship and goat rearing. Therefore, it is found that primarily 32% of the respondents have opined that they are doing retail business through their SHGs.

ANALYSIS OF COMPETENCY COMPONENTS

The spirit of entrepreneurship helps women to practice the art of innovating ideas, skills, competencies and other resources available to them. In this unit, the entrepreneurial competency of women is analyzed. The competency factors are duly correlated with personal traits and reasons stated by them for joining SHGs.

From the table4, it is clearly inferred that the respondents have primarily opined that they do not mind trying alternative ways when they meet failures. Followed by it, the second most important factor is that they are optimistic of their achievement in the enterprise.

In the next order, the respondents have expressed that they take people into confidence at the time of assigning task. In the fourth order, the respondents have viewed that they weigh pros and cons before taking decisions.

The respondents have strong conviction to succeed in their ventures is ranked the fifth order. In the sixth order, they have ranked the confident of solving any problem that come across in their business. Again the respondents have said that they are willing to accept challenges in new ventures.

The respondents further opined that they are capable of converting adversaries into opportunities. Entrepreneurship requires assuming risk of any kind is ranked in the ninth order. Finally, the respondents have expressed that they will not take rest until reaching the goal.

Therefore, it is concluded that the respondents have primarily opined that they do not mind trying alternative ways at the time of failures.

Table-4: Respondents Level of Perceptions on Entrepreneurial Competency

Sl. No	Statements	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Total Score	Total Mean	Rank
1.	I am willing to accept challenges in new ventures	176 (44.00)	210 (52.50)	14 (3.50)	0	0	638	1.60	7
2.	I am capable of converting adversaries into opportunities	186 (46.50)	193 (48.25)	20 (5.00)	1 (0.25)	0	636	1.59	8
3.	Entrepreneurship requires assuming risk of any kind	197 (49.25)	180 (45.00)	22 (5.50)	1 (0.25)	0	627	1.57	9
4.	I do not mind trying alternate ways When I meet failures	129 (32.25)	181 (45.25)	36 (9.00)	50 (12.50)	4 (1.00)	819	2.05	1
5.	I have strong conviction to succeed in my ventures	152 (38.00)	219 (54.75)	27 (6.75)	2 (0.50)	0	679	1.70	5
6.	I am confident of solving any problem that I may come across in my business	189 (47.25)	189 (47.25)	18 (4.50)	3 (0.75)	1 (0.25)	645	1.61	6
7.	I am optimistic about my achievement in my enterprise	136 (34.00)	225 (56.25)	38 (9.50)	1 (0.25)	0	722	1.81	2
8.	I take people in to confidence when I assign task	128 (32.00)	244 (61.00)	26 (6.50)	2 (0.50)	0	702	1.76	3
9.	I weigh pros and cons before I take decisions	159 (39.75)	225 (56.25)	14 (3.50)	2 (0.50)	0	693	1.73	4
10.	I will not rest until I reach the goal	227 (56.75)	145 (36.25)	26 (6.50)	2 (0.50)	0	603	1.51	10

Sources: Primary Data.

Factors Considered as Significant for Entrepreneurial Competency

Factor analysis is a multivariate statistical technique used to condense and simplify the set of large number of variables to smaller number of variables called factors.

This technique is helpful to identify the underlying factors that determine the relationship between the observed variables and provides an empirical classification scheme of clustering of variables into groups called factors.

Factor Analysis

Using all ten factors considered as entrepreneurial competency of Women started from willingness to accept challenges in new venture X1, X2,……, X10 – aim to reach goal, factor analysis is performed and results are presented in following tables:

Table 5 gives the rotated factor loadings, communalities, Eigen values and the percentage of variance explained by the factors.

Out of the 10 statements, 5 factors have been extracted and these five factors put together explain the total variance of this variable to the extent of 64.09 per cent.

In order to reduce the number of factors and enhance the interpretability, the factors are rotated. The rotation increases the quality of interpretation of the factors. There are several methods of initial factor matrix to attain simple structure of the data.

Table-5: Rotated Factor Loadigs

Variable	Factor					
	1	2	3	4	5	
X1: Willing to accept challenges in new ventures.	.552	-.297	-.203	.208	-.079	0.484
X2: Capable of converting adversaries into opportunities.	.240	.416	-.292	-.438	.596	0.864
X3: Entrepreneurship requires assuming risk of any kind.	.411	.574	.014	-.326	-.294	0.691
X4:T.rying alternate Ways When I meet failures.	.497	-.631	-.029	-.071	-.179	0.683
X5:Strong conviction to succeed in my ventures.	.593	-.159	.180	-.491	-.175	0.68
X6: Confident of solving any problem that I may come across in my business.	.304	.429	-.207	.512	-.008	0.582
X7: Optimistic about the achievement in the enterprise.	.373	.027	.504	.226	.370	0.581
X8: Take people in to confidence when I assign task.	.482	-.285	-.359	.176	.372	0.612
X9: Weigh pros and cons before I take decisions.	.046	-.097	.712	.005	.259	0.586
X10: Will not rest until I reach the goal.	.473	.501	.211	.266	-.237	0.646
Eigen Value	1.588	1.389	1.227	1.124	1.081	6.409
per cent of var. explained	15.879	13.893	12.271	11.240	10.810	
Cum. per cent explained	15.87	29.771	42.042	53.282	64.092	

Sources: Primary Data.

Table-6: Clustering of Inducing Variables into Factors

Factor	Inducing Variables	Rotated Factor Loadings
Opportunity Identification I. (15.879 %)	X1: Willing to accept challenges in new ventures.	0.826
	X2: Capable of converting adversaries into opportunities.	1.058
Risk Taking II. (13.893%)	X3: Entrepreneurship requires assuming risk of any kind.	0.984
	X4:T.rying alternate Ways When I meet failures.	1.126
Problem Solving III. (12.271 %)	X5:Strong conviction to succeed in my ventures.	0.563
	X6: Confident of solving any problem that I may come across in my business.	0.71
Systematic Planning & Decision Making IV. (11.240 %)	X7: Optimistic about the achievement in the enterprise.	1.164
	X8: Take people in to confidence when I assign task.	0.452
Persistence & Persuasion V. (10.810%)	X9: Weigh pros and cons before I take decisions.	0.942
	X10: Will not rest until I reach the goal.	1.109

Sources: Primary Data.

Five factors were identified as being maximum percentage variance accounted. The two variables X1and X2 were grouped together as factor I and accounts for 15.879% of the total variance. Factor I is named as Opportunity Identification,

A mindset is where one is trained to look for business opportunities from everyday experiences. The two variables X3 and X5 constituted the factor II and account 13.893% of the total variance, and is named as Risk Taking: Ability to assume any type of risk involved in business as well as the conviction to convert it in to opportunity. Variables X4 and X6 constituted the factor III, which is named as Problem solving i.e., observing the symptoms, diagnosing and curing and accounts 12.271% of the total variance.

The two variables X7 and X9 constitute the factor IV and accounts for 11.240% of the total variance and is named as Systematic Planning & Decision making: Breaking up the complex whole into parts, close examination of the parts and inferring about the whole; e.g. simultaneously attending to production, marketing and financial aspects (parts) of the overall business strategy (the whole).

Rest of the variables X8 and X10 constitute the factor V and named as Persistence & Persuasion - A 'never say die' attitude, not giving up easily, striving information seeking continuously until success is achieved. Eliciting support of others in the venture, this accounts for 10.810% of the total variance.

The ten factor analysis condensed and simplified the 10 inducing variables and grouped into five factors explaining 64.09% of the variability of all the ten inducing variables.

FINDINGS

Socio economic and demographic factors are often seen as critical factors influencing the involvement of women in managing their enterprise. Researchers have indicated that some degree of association can be established between entrepreneurial characteristics and performance.

D.D. Sharma et al., had expressed that socio- economic factors do also play a significant role in developing potential Entrepreneurs. From the data analysis it is found that 49% of the respondents belong to the age group of 30-40 years. This age group represents the economically productive women folk in the study area.

Distribution based on educational qualification shows that 67% of the respondents have opined that they have consummate school level education followed by those with either a diploma level education or informal education.

Thus, it is inferred that majority of the respondents have basic literacy. Majority of the respondents are married with their spouses in to entrepreneurial ventures. Among the sample population surveyed, 35.92% of the respondent's family profession is business and 13.06% are associated with other occupations like seasonal businesses.

Nearly 7% of the respondents' family members are involved in un-organized business sectors like vegetable vending, florists, livestock maintenance etc. Rest of them is employed in various areas like bakeries, power looms, weaving, and working as Artisans. Since majority of them are with previous entrepreneurial experience they find it easy to start a business as soon as they form their SHGs.

Their entrepreneurial competency level knobs them perform better overcoming all sorts of personal and professional hurdles. Based on the nature of business carried by the SHGs of the respondents, they are distributed as, 32% in retail business, 28.75% in service activities, 19.75% of them in manufacturing business and rest of them are involved in wholesale business or involved in other allied sector activities like fishery, dairy, piggery, poultry, cow, ship and goat rearing.

Factors like initial investment, business risk and knowledge and kelp available in marketing activities, mainly decide the choice of business run by SHGs. Majority of them are in retail business since this involves less managerial activities, they can develop their business based on their neighborhood networks , credit facilities offered by wholesale suppliers and comparatively low risk business.

Regarding the response to their rating of entrepreneurial competency components the respondents have top rated the statement that they do not mind trying alternative ways when they meet failures. Optimistic of their achievement in the enterprise is the next rated statement.

In the next order, the respondents have expressed that they take people into confidence at the time of assigning task followed by the habit of weighing pros and cons before taking decisions. Strong conviction to succeed in their ventures is ranked as fifth and problem solving skills catches the next order. Willing to accept challenges in new ventures and capability of converting adversaries into opportunities are the next two ranks. Facing and managing risk and perseverance are the last two ratings.

Out of the 10 statements framed to identify the entrepreneurial competency components, five factors were identified as being maximum percentage variance accounted. Factor analysis was carried out to derive out the entrepreneurial competency components.

Out of ten Statements framed for analyzing the opinion of respondents regarding entrepreneurial competency, the following factors are derived out factor analysis. Factor I (variables X1and X2) accounts 15.8795 of the total variance is named as **Opportunity Identification,** a mindset where one is trained to look for business opportunities from everyday experiences. Factor II (variables X3 and X5) accounts for 13.893% of the total variance, and is named as **Risk Taking. V**ariables X4 and X6 constituted the factor III, which named as **Problem solving** and accounts for 12.271% of the total variance. Factor IV (variables X7 and X9) accounts for 11.24% of the total variance and named as **Systematic Planning & Decision making** Rest of the variables X8 and X10 constitute the factor V and named as **Persistence & Persuasion,** which accounts for 10.81% of the total variance.

CONCLUSIONS

Entrepreneurial competency level of women entrepreneurs depends on skills like Opportunity Identification, risk taking, problem solving and decision making skills. Entrepreneurial inputs given through training shall concentrate on imbibing these skill sets among SHG women entrepreneurs, which would determine their Entrepreneurial success.

Developing social networks can be considered as a suitable strategy, the culture of sharing and exchanging experiences, views and innovative ideas will knob them to face the challenges of bigger players.

Reorienting policies to focus on developing entrepreneurship among SHGs and providing support to improve their performance would be of vital necessity.

REFERENCES

[1] Anjugam, M., and Anita, S., (2007). Functioning and Performance of SHGs in Tamil Nadu, India: Economic Empowerment of Women, New Delhi: New Century Publications, pp. 189-201.

[2] Azad, Gulab Singh, (1989). Development of Entrepreneurship among Indian Women, SEDME, Volume: XVI, Number: 3.

[3] Gupta, C. B., and Khanka, S. S., (2003). Entrepreneurship and Small Business Management, New Delhi: Sultan Chand & Sons, 4th Edition.

[4] Ivy, Jeno S., (2007). Empowerment of Women / through SHGs", India: Economic Empowerment of Women, New Delhi: New Century Publications, pp. 158-163.

[5] Janaki, Radha Krishnan, (2007), Women Empowerment through Self- Help Groups, India: Economic Empowerment of Women, New Delhi: New Century Publications, pp. 189-201, 238-241.

[6] Kavitha, K. S., and Vasudevan, V. P., (2007). SHGs for the Success of Women Entrepreneurs, India: Economic Empowerment of Women, New Delhi: New Century Publications, New Delhi, pp. 129-134.

[7] Krishnamurthi, N. A., and Suresh, K. M., (2007). Role of SHGs in Women Empowerment, India: Economic Empowerment of Women, New Delhi: New Century Publications, pp. 48-54.

[8] Manimala, M. J., (2005). Entrepreneurial Heuristics' in Entrepreneurship Theory at the Crossroads: Paradigms and Praxis, 2nd Edition, New Delhi: Biztantra Publishing.

[9] Neera, Burra; Joy, Deshmukh-Ranadive, and Ranjani, K. Murthy, (2005), Micro-Credit, Poverty and Empowerment: Linking the triad' Sage Publications.

[10] Ramasamy, H., (1995). Productivity in the Age of Competitiveness: Focus on Manufacturing in Singapore. In Productivity in the Age of Competitiveness. APO Monograph Series 16, Asian Productivity Organization.

[11] Thomas, W.; Y., Man, and Theresa, Lau, (2000), "Entrepreneurial Competencies of SME Owner / Managers in the Hong Kong Services Sector: A Qualitative Analysis", Journal of Enterprising Culture, Volume: 8, Number: 3, pp. 235-254.

[12] Thomas, J. J., Surviving in the City; The Urban Informal Sector in Latin America, CT: Pluto Press, East Haven.

[13] Thompson, J. E.; Stuart, R., and Lindsay, P. R., (1996). "The Competence of Top Team Members: A Framework for Successful Performance", Journal of Managerial Psychology, 113, pp. 48-66.

*****

FOR ANY CLARIFICATION OR SUGGESTION, WRITE US:

Editor-In-Chief
Pezzottaite Journals,
24, Saraswati Lane, Bohri, Near Modern Dewan Beverages,
Jammu Tawi – 180002, Jammu and Kashmir, India.
(Mobile): +91-09419216270 – 71
editorinchief@pezzottaitejournals.net,contactus@pezzottaitejournals.net

IMPACT OF EMOTIONAL INTELLIGENCE ON EXECUTIVES PERFORMANCE IN FERTILIZERS INDUSTRY OF INDIA

V. Kalaiarasi[120] V. Vishnupriya Dharshini [121]

ABSTRACT

This research paper aims to investigate "The Impact of Emotional Intelligence on the Managerial Performance". The study is descriptive in nature which gives an overall picture of the state of affairs at a given point of time. Convenient sampling method is adopted for the selection of the sample size. A sample size of 140 employees was administered to find out the impact of emotional intelligence on managerial performance. The sampling unit is the Supervisory Level employees of Fertilizers Industry, India. The primary data was collected through mail questionnaires. Secondary data were also obtained from the Magazines and Journals. Simple percentage and regression analysis are the tools used to analyze the data. There is a relationship between emotional intelligence and overall performance of the supervisory level employees. The study found that employees with high emotional intelligence were associated with high levels of performance.

KEYWORDS

Emotional Intelligence, Fertilizers Industry, Executives Performance etc.

INTRODUCTION

Emotions refer to a feeling state (including physiological responses and cognitions) that conveys information about relationships. Emotions are intense feelings that are directed towards someone or something, and are considered to be critical factors in employee be behavior – Stephen P.Robbins.

Emotional intelligence is defined by Goleman (1998) as the capacity for recognizing our own feelings and those of others, for motivating ourselves, and for managing emotions well in us and in our relationships. Essentially there are four competencies:

- Understanding yourself, or self awareness.
- Managing yourself or self management.
- Understanding others, or social awareness.
- Managing others, or social skills.

Some argue that EQ is more important in the work place than IQ. In the work place, there are constant interactions which are occurring among the people who work there. While some of these interactions are positive, others are negative.

The effect that these interactions have on the company will also have an effect on the company's ability to remain competitive in its given market or industry. Having said that, the key fact that must be considered in order to make these interactions more positive is emotional intelligence. Humans are emotional creatures, and this is the first key towards understanding EI.
In the past, companies and educational institutions paid an undue amount of attention to one's IQ, or their general intelligence quotient. While IQ is important to some extent, few corporations take the time to measure one's emotional intelligence.

EI is equal or greater in importance to IQ simply because people who work for organizations must be expected to get along with the people they work with. If they are managers, they must be effective in working with and managing the employees.

REVIEW OF LITERATURE

Emotional intelligence appears to be catalyst for articulating vision and developing constructive relationships with organizational members. A recent study reported that untrained teams with high emotional intelligence members' performed as well as trained teams with low emotional intelligence members **Ashkanasy & Daus, 2002.**

According to Bunker (1997) leaders' first need to be aware of and to manage their own feelings of anxiety and uncertainty. Then they need to be aware of the emotional reactions of other organizational members and act to help people cope with those reactions. At the same time in this process of coping effectively with massive change, other members of the organization need to be actively involved in monitoring and managing their emotional reactions and those of others.

According to **Cherniss & Goleman (2002)** emotional intelligence is important at for success at work.

Emotional intelligence emphasizes understanding one's self and others **Cooper & Sawaf 1996**people with a high level of emotional intelligence can make a crucial difference in the following areas of management:

- Decision making,
- Leadership,
- Trusting relationships& teamwork,
- Customer loyalty & Creativity/innovation.
-

Craig R Seal & Andrianna Andrews-Brown(2010), have proposed a new theoretical and empirical model of emotional intelligence(EI). They proposed that emotional ability (potential capacity) has a moderating effect on the mediated relationship of emotional quotient(preferred patterns) and emotional competence(actual behaviors) on performance outcomes.

Goleman (1998) suggests in his book on emotional intelligence that emotionally intelligent people can effectively handle conflicts with organizational members.

Daniel Goleman (2001-2003) suggested that Emotional intelligence can be learned. For businesses of all kinds, the fact that emotional competencies can be assessed and improved suggests another area in which performance – and so competitiveness – can be upgraded.

[120] Assistant Professor, Sri Ramakrishna Institute of Technology, Tamil Nadu, India, kalaisrit@gmail.com
[121] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, prdharshini31@gmail.com

Gardner (2010) defines interpersonal intelligence as the ability to understand others, how they work, what motivates them and how to work cooperatively with people. Intrapersonal intelligence is the ability to understand oneself to form an accurate model of one self and use this model operates effectively in life. A combination of interpersonal and intrapersonal skills can be seen as emotional intelligence responds to them appropriately.

Hassan Jorfi, Saeid Jorfi, Sirous Korahi Moghadam (2010) indicates that emotional intelligence has a Positive impact on the performance level of the managers and employees.

Emotional intelligence is one of the few key characteristics that give rise to strategic leaders in organizations. **Hill, C., and Jones, G., 2001.**

According to **LeRoux and De Klerk 2001** interpersonal intelligence and intrapersonal intelligence are the main components of emotional intelligence

Luthans (2002) listed emotional intelligence as one of five constructs capable of being developed and managed for improved performance in organizations.

Mayer & Salovey (1997) suggested that emotional intelligence can vary to fit with specific personality traits, implying a facilitating (i.e., Moderator) effect on the effectiveness of personality in the prediction of work outcomes.

Neal, M. Ashkanasy, & Catherine, S. Daus, (2002) studied that the emotions in organizational settings has provided new and important insights into the way in which people in the organizations behave.

Praveen M., Kulkarni B. Janakiram & D.N.S. Kumar (2009) studied that emotional intelligence has an impact on the performance level of the managers and supervisors.

The study conducted by **Rahim and Psenicka (2002)** in seven countries, their study revealed that motivation, one of the dimension of emotional intelligence, is positively related with effective problem solving style.

Salovey & Mayer (1990) think of as EI will not necessarily be more effective than other managers in dealing with conflict among their employees. However, they will be able to learn to use conflict management skills more readily than will individuals who bring less EI to the job.

According to **Sternberg (1996)** social intelligence is highly valued in the workplace and contributes to effective management, because social intelligence allows

Susan Tee Suan Chin, R.N. Anantharaman and David Yoon Kin Tong (2011) studied that employees in the SMEs do experience some kind of experiences that affect their working attitudes. The initial study provides many insights especially to the areas where respondents show a low level of emotional intelligence. The low level of emotional intelligence is related to the nature of work involved especially when it involves a high level of stress

The roots of emotional intelligence are embedded in the intelligence testing movement of **Thorndike (1920)** who was one of the first psychologists to identify the aspect of emotional intelligence (EI) which he called as Social intelligence. He included it in many types of intelligences that people possess.

Zubin R. Mulla (2010) analysed that EI did not show significant impact on job performance, for individuals having high interpersonal interaction on their jobs; EI was significantly related to job performance. On the other hand, for individuals having low interpersonal interaction on their jobs, EI was not related to job performance.

RESEARCH METHODOLOGY

Research Design : Descriptive research

Research Instrument: Structured Questionnaire

Sampling Plan

- Sample Method: Non-Probability Sampling & Convenience Sampling.
- Sample Size: 50.
- Sample Unit: Employees who do not hold a supervisory position.

Sampling Design

Convenience sampling, as the name implies, is based on the convenience of the researcher who is to select a sample. Respondents in the sample are included in it merely on of their being available on the spot where the survey was in progress.

Source of Data

- Primary Data: Structured Questionnaire,
- Secondary Data: Journals, Booklets, & Company Data, etc.

RESULTS AND DISCUSSIONS

Table-1: Profile of the Respondents

	Category	Frequency	Percentage
Gender	Male	130	93
	Female	10	7
	Total	**140**	**100**
Age	21-30	4	3
	31-40	15	11
	41-50	73	52
	51 & Above	48	34
	Total	**140**	**100**
Years of Service	1-10	21	15
	11-20	31	22
	21-30	67	48
	31 & above	21	15
	Total	**140**	**100**
Educational Qualifications	Professional Qualification	56	40
	Other Qualification	84	60
	Total	**140**	**100**

Sources: Primary Data.

Table-2: Emotional Quotients of Respondents

S.No.	Particulars	Number of Respondents	Percentage
1	Extremely High	0	0
2	High	27	19
3	Moderate	91	65
4	Low	14	10
5	Extremely Low	8	6
	Total	140	100

Sources: Primary Data.

Table-3: Self Assessment of Respondents Performance at Work

S.No.	Particulars	Number of Respondents	Percentage
1	Extremely High	23	16
2	High	61	44
3	Moderate	38	27
4	Low	13	9
5	Extremely Low	5	4
	Total	140	100

Sources: Primary Data.

Regression Analysis

Table-4: Impact of Emotional Intelligence on Work Performance Based on Self Assessment by Respondents Coefficients*

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig
	B	Std.Error	Beta		
1 (Constant)	.230	.240		.957	.340
Self Awareness	.346	.106	.315	**3.258**	.001
Self Management	-.082	.158	-.057	-.516	.606
Social Awareness	.029	.115	.028	.252	.801
Relationship Management	.574	.162	.454	**3.549**	.001

Sources: Primary Data.

a. dependent variable: self assessment

Table-4 it is inferred that during self assessment by respondents, the factors self awareness and relationship management have great impact on the performance of the respondents.

The 't' value for self awareness is 3.258 and the 't' value for relationship management is 3.549 which prove that it has an impact on the performance.

The other factors namely self management and social awareness do not have any significant impact on the performance of the respondents.

Table-5: Impact of Emotional Intelligence on Overall Work Performance of Respondents

Coefficients*

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig
	B	Std.Error	Beta		
1 (Constant)	.230	.240		.957	.340
Self Awareness	.346	.106	.315	**3.258**	.001
Self Management	-.082	.158	-.057	-.516	.606
Social Awareness	.029	.115	.028	.252	.801
Relationship Management	.574	.162	.454	**3.549**	.001

Sources: Primary Data.

a. Dependent variable : overall performance

From table.5 it is inferred that the factor namely relationship management has greatest impact on the overall performance of the respondents. The t value for relationship management is 5.062 which prove that it has strong impact on performance. The other factors namely self awareness, self management and social awareness do not have significant impact on the overall performance of the respondents.

CONCLUSIONS

In this study confirms the premises that managers should have emotional intelligence to increase their performance. Managers need to discover the value of emotional intelligence to handle challenging situations at the organization. Further the study would give way for further research on the same or related issues.

REFERENCES

[1] Ashkanasy, N. M., Aand Aus, C. S.m (2002). 'Emotion in the Workplace: The New Challenge for Managers", Academy of Management Executive, 16, pp. 76-86.

[2] Bunker, K. A., (1997). 'The Power of Vulnerability in Contemporary Leadership", Consulting Psychology Journal, 49(2), pp. 122-136.

[3] Cherniss, C., and Goleman, D., (Eds). (2002). The Emotionally Intelligent Workplace. San Francisco: Jossey-Bass.

[4] Cooper, R. K., and Sawaf, A., (1996). Executive EQ Emotional Intelligence in Leadership and Organizations, New York: Perigee Books.

[5] Craig, R. Seal, and Adrianna, Andrews-Brown, (2010), An Integrative Model of Emotional Intelligence: Emotional Ability As A Moderator of The Mediated Relationship of Emotional Quotient and Emotional Competence, Organization Management Journal, pp-143-152.

[6] Gardner, J. K., and Qualter, P., (2010). "Concurrent and Incremental Validity of 3 Trait Emotional Intelligence Measures", Australian Journal of Psychology, **62**: 5–12.

[7] Goleman, D., (1998). Working with Emotional Intelligence, New-York: Bantam.

[8] Daniel, Goleman, (2001-2003), Working with Emotional Intelligence. Bantam Books.

[9] Hassan, Jorfi; Saeid, Jorfi, and Sirous, Korahi Moghadam (2010), "Impact of Emotional Intelligence on Performance of Employees", Postmodern Openings, Year: 1, Vol: 4, pp: 63-74.

[10] Hill, L., (2004), "New Manager Development for the 21st Century", Academy of Management Executive, 18, (3), pp. 121-126.

[11] Johnson, P. R., and Indvik, J., (1999). "Organizational Benefits of Having Emotionally Intelligence Managers and Employees", Journal of Workplace Learning, 11, pp.84-88.

[12] Le, Roux R., and De, Klerk R., (2001). Emotional Intelligence Workbook, Human and Rousseau (Pty) Ltd., Cape Town

[13] Luthans, F., (2002). "Positive Organizational Behaviour: Developing & Managing Psychological Strengths for Performance Improvement", Academy of Management Executive, Volume: 16, Number: 1, pp. 57-76.

[14] Mayer, J. D., and Salovey, P., (1997). What is Emotional Intelligence?, In: P. Salovey, & D. Sluyter, Emotional Development And Emotional Intelligence: Educational Implications, New York: Basic Books.

[15] Neal, M. Ashkanasy, and Catherine, S. Daus, (2002). "Emotion in Work Place: The New Challenge for Managers", Academy of Management Executive, Volume: 6, Number: 1, pp. 76-86.

[16] Praveen, M.; Kulkarni, B. Janakiram, and D., N. S. Kumar, (2009) "Emotional Intelligence and Employee Performanceas an Indicator for Promotion: A Study of Automobile Industry in India", International Journal of Business and Management, Volume: 4, Number; 4, pp161-170.

[17] Rahim, M., and Psenicka, C., (2002). "A Model of Emotional Intelligence and Conflict Management Strategies: A Study in Seven Countries", The International Journal of Organizational Analysis, 10(4), pp. 302-306.

[18] Salovey, P., and Mayer, J. D., (1990). "Emotional Intelligence", Imagination, Cognition and Personality, 9(3), pp. 185-211.

[19] Salovey, P.; Mayer, J. D.; Caruso, D., and Lopes, P. N., (2001). "Measuring Emotional Intelligence as a Set of Mental Abilities with the MSCEIT", In: S. J. Lopez, and C. R. Snyder (Eds.), Handbook of positive psychology assessment. Washington DC: American Psychological Association.

[20] Sternberg, R., (1996). "Successful Intelligence", NewYork: Simon & Schuster.

[21] Susan, Tee Suan Chin; R. N. Anantharaman, and David, Yoon Kin Tong, (2011). "Analysis of Level of Emotional Intelligence among Executives in Small and Medium Sized Enterprises", Journal of Human Resources Management Research, Volume: 2011 (2011),

[22] Thorndike, E. L., (1920). "Intelligence and its Use", Harper's Magazine, 140, pp. 227-235.

[23] Zubin, R. Mulla, (2010). Do Emotionally Intelligent People do Well in All Jobs? Exploring the Moderating Role of Inter Personal Interaction", Vision - The Journal of Business Perspective, Volume: 14, Number: 4, pp. 247-254.

*****

GLOBAL OPTIMIZATION OF PRODUCTION AND DISTRIBUTION PLANNING

M.Mohan[122] Kannan Jegathala Krishnan [123]
R.Niruban Chakravarthi[124]

ABSTRACT

Production and distribution planning is an absolute necessity for every industry to thrive in the highly competitive environment. Since the available resources for production such as raw materials, labour, funds, etc are limited, it is necessary to allocate them effectively and utilize the resources optimally to remain agile in the global market. Production plan can be classified into three types, namely short range plan, medium range plan and long range plan.

Production and Distribution Planning, a tool for tactical planning, is designed to handle the most complex optimization tasks related to finding best balance between the service level provided to customers and the costs of supply chain management (inventory level at stores, production costs, transportation costs, etc.).Production and Distribution Planning is successfully used in various industries such as the following: food, distribution, pharmaceutics, package printing, automobile production, engineering, metallurgy, electronics, etc.Simulated annealing (SA) is a generic probabilistic metaheuristic for the global optimization problem of applied mathematics, namely locating a good approximation to the global optimum of a given function in a large search space. It is often used when the search space is discrete (e.g., all tours that visit a given set of cities).

KEYWORDS

Production and Distribution Planning etc

INTRODUCTION

Usually, the production planning rarely manages the costs of distribution. When it makes it, these costs are supposed proportional to the quantities and\or to the distances and do not take into account the structure of the trips of delivery in less than truckload situation. As for the literature on the periodic routing, it supposes known the quantities to be delivered for every period and it thus intervenes on an already calculated and imposed plan of production. Now with Extension of the activities of companies, we must take in count different level of the supply chain: supply *(approvisionnement)*, production and distribution.The system efficiency can be improved by exploiting scale economies due to production/distribution coordination, by balancing production lots and vehicle loads, by reducing total inventories and rationally planning vehicle routes both in topological and temporal sense. The vehicle and inventory routing problem has been widely investigated, both under deterministic and stochastic assumptions, but usually ignoring production decisions [3, 9, 10, 11, and 13]. In contrast, the problem of production planning is widely Investigated in several details – lot sizing, scheduling,…but ignoring distribution decisions too. Several authors stressed the importance of an integrated approach of management while going from the first phases of the process to the final customer, they put in evidence a potential economic benefit discounted this integrated approach. Our work joins in an initiative of generalization in which we wish to build simultaneously a production planning and the trips of distribution, over several periods, to minimize the total cost. This approach is possible and interesting in sectors for rather regular demands and when production costs are of the same order as the costs of transport, as the cattle foods. We consider a problem where a plant makes a single product to feed n customers on a horizon of planning of T periods or "days". Every customer has a known geographic localization and a known consumption in advance throughout the horizon of planning. It has a zone of stocking of definite capacity and has to undergo no break. The daily capacity of production of the plant is limited. The product can be stored to the plant, with a limited capacity, or be elivered to the customers by trips of vehicles. To do it, we have a fleet of K vehicles with homogeneous capacity W based at the plant. We suppose that every vehicle makes in most one trip by period and that every customer can be delivered for most one time by period. We consider a cost of setup for every day of production, a cost of stocking to the plant and costs of transport proportional to the distance, but no costs of stocking at the customers. The objective is to determine the days of production; the quantity of product delivered to each customer every day and the associated trips, in order to minimize the total cost. This problem is NP-hard because it comes down to the vehicle routing problem in the monoperiod case. To approach well this problem, we shall begin with a review of literature, followed by a modelling of our problem in section 3, in section 4; we shall see the heuristics of resolution which will be followed by the results in section 5 and the conclusion in the section 6.

LITERATURE REVIEW

We note that the literature approaching the problem of the coordination between the planning of the production and the distribution is not plentiful. A model of coordination between the production planning and transport was developed by Dror and Ball [5]. It considers a plant producing only one product and supplying customer over only one period. This model is simplistic because to enlarge the analysis to multiperiod horizon, recursive solutions of single period models or corrections to cost coefficients have been proposed. Chandra and Fisher [1] treated this problem by considering a multiproduct case, but their approach is more directed to the organization and scheduling of the production seen the multiproducts case and its combination with the distribution. They developed a heuristics which has allowed a reduction of costs varying between 3 % and 20 % compared with an uncoupled approach. Another study of Chandra [2], which is this time directed much more to the planning and to the dynamic case, it concerns the launch of the commands from a regional platform of storage to satisfy the demands of the customers of the same region. The tests made on 33 different

[122] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, mmohan1986@gmail.com
[123] PG Scholar, School of Engineering & Science, Victoria University, Australia, Kannan.Jegathala-Krishnan@vu.edu.au
[124]Lecturer, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, niruban.leo@gmail.com

cases showed a reduction of cost which varies between 5 % and 14 % stemming from the coordination between the distribution and the launch of the commands. Fumero and Vercellis [8] also dealt with the problem of coordination between the production and distribution. They were interested in the configuration with multi products and multi periods with a vision directed more towards scheduling concerning the production part, they speak about synchronization between the production and the distribution. A Lagrange relaxation was used to solve this problem; it was tested on Instances with maximum of 12 customers, 10 products and 8 periods. The results of the Tests were compared with an uncoupled approach and showed a significant economy compared to the results of the uncoupled approach. Zipkin [14] treated the problem of coordination between the production and the Distribution between several factories. A lower bound was found for this problem where the finished products of some factories will serve for the production of the other factories not being on the same site. Ertogral and al. [7] handled a similar problem; their approach consisted in decomposing the global problem and resolving a problem of planning of production and a problem of distribution. Supplementary constraints will be afterward introduced to insure the coordination and adapt the results of the first phase to the global problem. The problem of coordination between the production and the distribution was also adapted in the field of the sorting and of the routing of the post. We find these works in Metter's [12] where the objective was to coordinate the organization of the activities of the sorting and that of the routing of the mail in the various centres of distribution to Reduce the cost and the delays of routing.

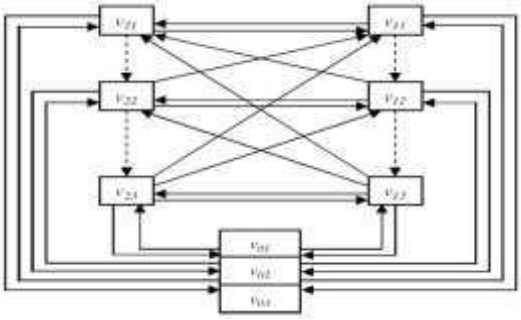
3. Model Formulation

Our problem is defined as follows.

Let G = (N, A) be a graph (spatiotemporal)

where N = {*v01*, *v02*,...,*v0T*, *v11*, *v12*,...,*v1T*, *v21*, *v22*,...., *v2T*,.....,*vn1*, *vn2*,...,*vnT*} is a set of vertices corresponding to demand of each customer, the vertex *v0t* correspond to factory through the horizon of planning. A = {(*vit*, *vjt'*): *vit*, *vjt'* ⊓¯N, *it* ⊓¯*jt'*} is the arc set. The cost of the segment *vit*, *vit'* is zero because it's the same customer and the cost is the inventory cost supported by the customer and not taken into account in our cost, but the cost of *vit*, *vjt* (or *vjt'*) has the cost *cij* of travel from the vertex *vit* to the vertex *vjt* (or *vjt'*). It's clear that the Transition from the vertex *vit* to the vertex *vit'* (the same customer) with *t'* ≺ *t* is prohibited because before satisfy the demand *qit* we must satisfy the demand *qit'* before (Not authorized break). This configuration is represented by the variable *r it jt X* (, ') . An example with two customers and three periods is given in the figure 1.

Figure-1: Model with 2 Customers and 3 Periods



Sources: Authors Compilation

→ Not null cost of transition

Variables and parameters of the model are defined as follows

Variables:

QPt: Quantity produced in the period t

QLt: Quantity delivered in the period t

Iit: level of the stock of the customer (i) in the period (t)

.ty = 1 if there is setup of production in the period t 0 if not

Parameters

N: set of n customers

T: horizon of planning

W: Capacity of vehicles

K: Total number of vehicles

r: index of trip

R: total number of trip

qit: demand of customer (i) in period (t)

CP: Production capacity

CSCi: Storage capacity of customer (i)

CSD: Storage capacity of the factory

c(it,jt'): Travel cost from the vertex (vit) to the vertex (vjt')

h: Cost of Unitarian stocking by period

Cl: Cost of setup of the production

$$\text{Min} \sum_{r=1}^{R} \sum_{v_{it} v_{jt'} \in V} C_{(it,jt')} X^{r}_{(it,jt')} + \sum_{t=1}^{T} y_t C_l + \sum_{t=1}^{T} h * I_t \qquad (1)$$

$$\sum_{r=1}^{R} \sum_{v_{jt'} \in V} X^{r}_{(it,jt')} = 1 \qquad v_{jt'} \in V \ (it \neq jt') \text{ and } (i \neq 0) \qquad (2)$$

$$\sum_{r=1}^{R} \sum_{v_{it} \in V} X^{r}_{(it,jt')} = 1 \qquad v_{it} \in V \ (it \neq jt') \text{ and } (i \neq 0) \qquad (3)$$

$$\sum_{r=1}^{R} \sum_{v_{it} \in V} X^{r}_{(it,jt')} - \sum_{r=1}^{R} \sum_{v_{jt'} \in V} X^{r}_{(jt',kt)} = 0 \qquad v_{it} \in V \qquad (4)$$

$$\sum_{r=1}^{R} \sum_{v_{it} \in V} X^{r}_{(it,jt')} = \sum_{r=1}^{R} \sum_{v_{jt'} \in V} X^{r}_{(jt',kt)} \qquad (5)$$

$$\sum_{r=1}^{R} \sum_{t=1}^{T} X^{r}_{(0t,it')} \leq K \qquad t=1..T \qquad (6)$$

$$\sum_{v_{it}, v_{jt'} \in V} X^{r}_{(it,jt')} q_{jt'} \leq W \qquad r=1..R \qquad (7)$$

$$\sum_{v_{it}, v_{jt'} \in S} X^{r}_{(it,jt')} \leq |S| - 1 \quad (S \subset V \mid |v_{0t}| \geq 2 \leq |S| \leq n) \ (v_{0t} \notin S) \text{ for each } r \qquad (8)$$

$$\sum_{r=1}^{T} X^{r}_{(it,jt')} q_{jt'} \leq CSCi - I_i \qquad i=1..T, i=1..n \text{ and } r=1..R \qquad (9)$$

$$QP_t \leq CP \qquad t=1..T \qquad (10)$$

$$QL_{0t} = \sum_{r=1}^{R} \sum_{t=1}^{T} \sum_{v_{it}, v_{jt'} \in V} X^{r}_{(it,jt')} q_{jt'} \qquad t=1..T \qquad (11)$$

$$QP_t \leq CSD - I_t \qquad t=1..T \qquad (12)$$

$$I_t = I_{t-1} + QP_t - QL_t \qquad t=1..T \qquad (13)$$

$$I_{it} = I_{it-1} + \sum_{r=1}^{T} X^{r}_{(it,jt')} q_{jt'} - q_{it} \qquad i=1..T, i=1..n \text{ and } r=1..R \qquad (14)$$

Our objective is to minimize the total cost of production (setups and inventory) and distribution; these costs are included in the objective function (1). Only one visit is authorized by vertex, except the vertex v0t which can be visited several times per period following the number of vehicles in service at this period, constraints (2) and (3) ensure this condition and constraints (4) that the vehicle how visit vertex leave it after delivery. Each trip begins and ends at the factory, constraint (5), and a constraint (6) ensures that number of vehicle is not exceeding K vehicles every period. The vehicle capacity restriction is imposed by constraint (7). The constraints (8) eliminate all subtours defined over subset N\{v0t}. Constraints (9) ensures that the storage capacity of customer are not Exceeded, the analogous restriction is ensured by constraints (10) and (11) for the capacity of production and storage of the plant. Constraints (12) determine the total quantity delivered in the period (t) to all the customers. Constraints (13) and (14) are

the inventory balance constraints for the factory and each customer respectively.

ALGORITHM

For the resolution of our problem, we developed a greedy heuristics called Coupled Heuristics. It bases itself on an (iterative) process consisted of 3 main phases: Determination of the quantities to deliver in every period and to every customer Construction of the trips of distribution (distribution plan). Determination of date of manufacturing of quantities to be delivered.

4.1 Determination of Quantities

From the current period t, we verify if all the demands of this period were satisfied. If it's true, there will be neither delivery nor production, and the cost of this period is zero.Otherwise a delivery must be made to avoid the breaks. The quantity to be delivered could be produced in the same period or draw from the stocks of the plant. To decide on the number of period to be satisfied from the current period, the unsatisfied Total demands by period are accumulated while being careful not to violate the constraints. We consider the capacity of production *CP* (constraint 10), if we decide to produce this quantity in the current period. We consider also the capacity of distribution *K*W* (constraints 6 and 7). Concerning the customers, we shall take into account the storage capacity *CSC* of each one of them (constraint 9). If one of customers can't receive his demand for the futures periods, we don't deliver it. Having to determine the number of periods to satisfy and the quantities to deliver for every customer, we pass then to the construction of the trips of distribution.

4.2 Construction of Trips of Distribution

For the construction of the trips of distribution, we have adapted the algorithm of Clarke and Wright for the vehicle routing [4] to our problem. Not to exceed the number *K* of available vehicles (constraint 6), we continue to make the fusions (mergers) of the trips even if the cost increases and it to has to reach the number of available vehicles, but we shall take the care of taking the feasible fusions (mergers) which respect the constraint capacity *W* of vehicle and which represents the weakest increase of cost of distribution for the current period. To improve still the cost of the distribution plan founded, a local search of type of the type two optimal is going to be applied on all the trips of this plan. This local search is applied to every trip and between trips, with first & better improvement.

4.3 Determination of Date of Manufacturing

To make the decisions of production/storage and elaborate the production planning, we have to adopt an approach inspired by the Wagner and within algorithm [6] for the optimal planning of the production, but as we have already underlined it, the coupled heuristics is Greedy and it proceeds period by period. Having to fix the quantity to produce and to elaborate the distribution plan for period t, our heuristics is going to fix the date of manufacturing this quantity. To be made, we are going to compare the cost of a setup of production with regard to a cost of stocking the quantity that to be delivered since the last period when there was setup of production. If the cost of storage is lower than the cost of setup of production and the capacity of production and storage is not violated, we are going to choose the stocking option, if not, a setup of production will take place. Having to fix the quantity to be made for the current period, its date of manufacturing and the plan of distribution, we pass in the next period and we unwind again the same process to have the end of the horizon of planning.

COMPUTATIONAL RESULTS

5.1 Heuristics of Comparison

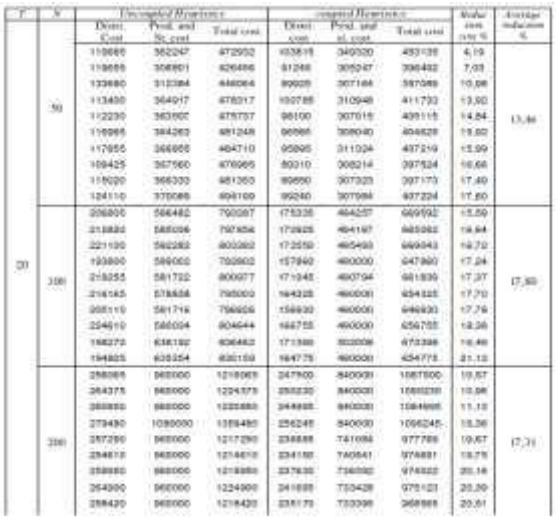
To compare our heuristics with the classic approach, we developed a constructive Heuristics based on an uncoupled approach. This heuristics is called an Uncoupled Heuristics. For this last one, we elaborate at first the production planning, without taking Into account the cost and configuration of future trips. Building us at first a production Planning with the Wagner and Witting method [6], this plan is going to be to imposed on The distribution. After that, a distribution planning is developed with a Clarke and Wright method [4] followed by a post optimization based on two optimal method like in coupled heuristics. The distribution is made in the first of horizon of planning, we must satisfy the demand of current period at first, and if we have a sufficient transportation capacity and a suffisante storage capacity at the customers, we shall pass to a satisfaction of the demands of the next periods to has the satuation of the storage capacity of the customers or the capacity of transportaion; or until the next setup of production because the inventory at the plant are Exhausted and the production planning are imposed to the distribution. We do this in order to benefit from the possibility of storing free of charge at the customer's. If the distribution is not made until the next setup of production, we must do it in the period when unsatisfieddemands exist. The same process is resumed (saturation of distribution capacity or storage Customers) or until the setup of production and it to has the end of the planning horizon.

5.2 Tests and Comparison

The both heuristics described in this paper were applied to 30 distinct test cases to determine the importance and the consequence of the coordination between distribution and production planning. The test cases were organized into three data sets. Data set 1 contained 50 customers (sur) 20 periods, data set 2 contained 100 customers (sur) 20 periods and data set 3 contained 200 customers (sur) 20 periods. Seen that there are no test cases known for this problem, the test cases were creating with an (unpredictable, random, aléatoire) generator. For each set of instances, we have adapted the parameters of both heuristics (suivant) the number of customers -Production capacity and storage capacity of palnt and customers, number of vehicles, and the cost of production setup-. To test the (interest interet) of the coordination between the production and the distribution, we implemented both coupled and uncoupled heuristics in Delphi 7 and we tested it on a PC 2.8 Ghz. Times of calculations are unimportant, (de l'ordre de la seconde pour les plus grandes instances) are of the order of second for the biggest instances for the both heuristics. We compared the results of coupled heuristics with an uncoupled heuristics. We compared both heuristics developed on the basis of the total cost (distribution, storage and production costs). To quantify the reduction of cost to bring by the coupled approach.For most approached our configuration of the problem, we have developed a test cases and parameters with cost of distribution which represents a not unimportant part of the global cost, it varies between 20% and 30% of the global

cost. The results founded are represented in the table 2. We notice at once that we have a significant reduction stemming from the coordination between the production planning and the construction of the trips of distribution. This reduction varies between 4,19% and 21,13% if we consider all the test case. If we look at the results in the detail, the set 1 contained 50 customers has an average reduction of total cost of 13,46%. We notice as well as that the coupled heuristics manage to make fall, for this set, the costs of the production-storage and the costs of distribution simultaneously. To note also that seven cases tests on ten are above the average reduction of total cost, what shows the efficiency of the coupled heuristics. Concerning the set 2 contained 100 customers, the average reduction cost is 17,80%. Three cases tests are above the average reduction of total cost, but the rest of the tests cases are not very remote from the average reduction of total cost, (2,21% for the worse test case). We also notice that the vast (l'étendu) of the reduction of total cost is smaller compared with the set 1. We also notice that the average reduction of total cost for this set is greater then avreage reduction of set 1. About set 3 contained 200 customers, the avreage redction is 17,31%. Seven tests cases are above the avreage reduction of total cost. The efficiency of coupled heutistics is proved for this set of test case too. Concerning the set 2 and 3, the reduction of total cost is greater then reduction of total cost of set 1.

Table-2: Summary Table of Results of Both Heuristics



Sources: Authors Compilation
***N*: Customers number;**
***T*: Number of period.**

CONCLUSIONS AND PERSPECTIVES

The problem aimed in this work is very complex because it integrates sub-problems which are already difficult. We believe that the model presented in this paper provides a practical effective approach to the coordination between the production and distribution planning. Although a simplified problem was studied at first, the results confirm the importance of a global optimization of the production and distribution planning. We have developed an effective heuristic for the integrated optimization of production and distribution planning. Tested on three groups of test case, she allows making significant reduction of total cost compared to an uncoupled approach. We work at present on more complex neighborhood modifying simultaneously days of production and trips of distribution. The purpose shall to be able to use a met heuristics to more improve the already found solutions. Other perspective of this work is to generalize the problem to the multiproduct case with costs of production and storage who can be different from one product to the other one.

REFERENCES

[1] P., Chandra, and M., L. Fisher, (1994). "Coordination of Production and Distribution Planning", European Journal of Operational Research, Volume: 72, Issue: 3, pp. 503-517.

[2] P., Chandra, (1993). "A Dynamic Distribution Model with Warehouse and Customer Replenishment Requirements, Journal of Operational Research Society, Volume: 44, Number: 7, pp. 681-692.

[3] M., Christiansen, and B., Nygreen, (1998). "A Method For Solving Ship Routing Problems With Inventory Constraints", Annuals of Operations Research, 81, pp. 357-378.

[4] G., Clarke, and J., W. Wright, (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points, Operations Research", 12, pp. 568-581.

[5] M., Dror, and M., Ball, (1987). "Inventory/Routing: Reduction from an Annual to a Short-period Problem", Naval Research Logistics, Volume: 34, pp. 891-905.

[6] L., Dupont, La gestion industrielle, (1998). Edition Hermes, pp. 171-173.

[7] K., Ertogral; S., D. Wu,, and L., I. Burke, "Coordination Production and Transportation Scheduling in the Supply Chain", Technical Report, Lehigh University, 1998.

[8] F., Fumero, and C., Vercellis, (1999). Synchronized Development of Production, Inventory, and Distribution Schedules", Transportation Science, Volume; 33, Number: 3, pp. 330-340.

[9] M., Gendreau; G., Laporte, and R., Séguin, (1996). "Stochastic Vehicle Routing", European Journal of Operational Resaerch, 88, pp. 3-12.

[10] G., Ghiani; F., Guerriero; G., Laporte, and R., Musmanno, (2003). "Real-time Vehicle Routing: Solution Concepts, Algorithms and Parallel Computing Strateges", European Journal of Operational Resaerch, 151, pp. 1-11.

[11] B., R. Golden, and E., A. Wasil, (1987). "Computerized Vehicle Routing in Soft Drink Industry", Operations Research, Volume: 35, Number: 1, pp. 6-17.

[12] R., D. Metters, (1996). Interdependent Transportation and Production Activity at United States Postal Service", Journal of Operational Research Society, 47, pp. 27-37.

[13] P., Toth, and D., Vigo, (2002). "Models, Relaxations and Exact Approches for the Capacitated Vehicle Routing Problem", Discrete Applied Mathematics, 123, pp. 487-512.

[14] P., Zipkin; J., Bramel, and S., Goyal, (2000) "Coordination of Production / Distribution Networks With Unbalanced Leadtimes", Operations Research, Volume: 48, Number: 4, pp. 570-577.

*****

PERFORMANCE STUDY ON INFLUENCE OF TOOL DESIGN ON ELECTROCHEMICAL MICROMACHINING

L. Balasubramanian[125] S. Dharmalingam[126]

ABSTRACT

Electrochemical Micromachining (EMM) is a promising technology for machining the microstructures and components. EMM appears to offer several advantages including better precision, wide range of materials could be machined and no stress on the machined workpiece. This paper is focused on development of an EMM system in which the micro structure and components are fabricated. Influences of tool electrode design (flat, conical with rounded and wedged electrode) on machining rate and overcut by varying the electrolyte concentration are discussed in detail.

The wedged electrode produces higher machining rate and conical with rounded electrode improves the micro-hole accuracy by 2.2 times compared to flat electrode. Finally, applications of an electrochemical micromachining are also discussed.

KEYWORDS

Electrode, Tool Design, Electro-Chemical Micro-Machining, Applications, Machining Set-Up etc

INTRODUCTION

The need for developing new, multi-functional and multi-material micro-components has considerably increased and improving manufacturing competency due to increased competition have posed new challenges in the manufacturing sectors (Qin et al. 2010).

Micromachining is the process of removing material in the form of chips or debris having size in the scale of micron. Among the various competent processes, EMM is considered for its following advantages such as no residual stress, versatility to machine any kind of material, no problem of heat affected zone, no tool wear, short machining time, cost effective, high precision can be achieved and the quality of good surface finish makes this process more attractive for drilling holes on the components exposed to high temperature (Bhattacharyya et al 2004).

EMM is a process of removing material similar to electroplating. The workpiece in EMM is dissolved according to Faradays laws of electrolysis. In this process, the workpiece to be machined is made as anode and the tool is made as cathode of an electrolytic cell with a salt solution being used as an electrolyte. On the application of a potential difference between the two electrodes with availability of adequate electrical energy between the tool and the workpiece, the positive metal ions leave the workpiece (Jain 2002).

[125]Assistant Professor, Kathir College of Engineering, Tamil Nadu, India, balakce@yahoo.com
[126]Associate Professor, Kathir College of Engineering, Tamil Nadu, India, dharmanit@gmail.com

In last decade EMM received much attention in the fabrication of micro-parts/ micro-components (Datta & Landolt. 2000). Keeping this in view, an indigenous development of EMM machine set-up has been considered to carry out a thorough research for achieving the satisfactory control on EMM process parameters to meet the micromachining requirements. The following Table 1 summarizes the developments in machine building worldwide.

Table-1: Developments in EMM Machine Set-Up

S.No.	Researchers & year	Purpose	Machine significance
1	Bhattacharyya et. Al & 2002	Micro-hole machining	Single axis (Z-axis), Microprocessor based IEG control, Feeding mechanism: Stepper motor and ball screw. Pulse power supply frequency in millisecond.
2	Yong et al & 2003	Micro-hole & 3-D Micro structures with high aspect ratio machining.	3 axis machine setup. Z axis with resolution of 0.1µm, XY stage with resolution of 0.1µm, Computer based closed loop IEG control, pulse power supply frequency in millisecond
3	Kozak et al & 2004	2 ½ D and 3D micro components	Drive System • Motor type: DC • Maximum travel :15mm • Resoltion :7nm • Number of axes:3 • Maximum velocity:1.4 mm/sec • Operating software: Win move Power supply • Operation mode: pulse or DC • Current output:0-1A • Voltage output:1-15V • Pulse frequency: 10Hz-1000KHz
4	Kurita et al & 2006	3D shape machining	3 axis machine setup. Z axis with resolution of 0.0156 mm/ pulse, XY stage actuated by piezo transducer, electrolyte supply with flow rate of 1-5 ml/min Computer based closed loop IEG control, pulse power supply frequency in millisecond.
5	Zhao et al & 2006	Deep hole drilling and milling	Multifunctional machine tool, 3 axis numerical controlled electrode feed mechanism, XY or Z axis is driven by servomotors

			through ball screw with a resolution of 0.1 µm. Electrodes rotates in the range of 1000 to 25000 rpm. High frequency pulse power supply
6	Lipiec et al & 2007	Micro-hole machining	The test stand consists of seven main parts • the machine body with fixed drives set with voltage pulse transmission system and the table for sample fixing, • observation system consisting of laboratory microscope of 100 times magnitude, • nanosecond pulse generator, • drive control system, • electrolyte delivery system, • PC computer with driving software, • ventilation system
7	Zhang and Zhu &2008	Micro-hole machining	Aaxis machining with high feed accuracy of 0.08 µm Motor drive, stepping motor, harmonic gear decelerator, ball screw and motion parts.IEG is controlled is maintained at 5 µm.
8	Tusi et al & 2008	Micro-hole machining	3 axes machine setup , with high speed spindle installed on z axis moving stage, pulsed power supply in microseconds.
9	Ryu & 2009	Micro-hole machining	3 axis machine setup, nano positioning system and controller, vision system, oscilloscope, High frequency pulse generator. High frequency pulse power supply.

Sources: Authors Compilation

Although the many researchers have developed the advanced EMM machine tool but still an in-depth research is required to overcome the practical problems in EMM tooling development. The tool shape is an exact negative mirror image of the workpiece to be produced. One of the main considerations for EMM tooling is the flow of the electrolyte solution. The tool design should provide not only the cathode dimensions but also an appropriate electrolyte path to prevent undesired overcut. Hence, the proper design of micro tool is required to eliminate the prevailing problems. The following Table 2 summarizes the developments in tool design worldwide.

Table-2: Developments in EMM Tool Design

S.No	Researchers & Year	Tool design used for EMM	Outcome
1	Ahn et al & 2004	Rounded Electrode	Improves mach inability
2	Westley et al & 2004	Curvature Shape Tool Electrode	Improves electrolyte flow
3	Bhattacharyya et al & 2004	Vibrating Tool Actuated By Piezo Transducer (PZT)	Improves electrolyte flow, MRR and accuracy
4	Ebeid et al & 2004	Low Frequency Vibrating Tool	Improves the accuracy
5	Lee et al & 2007	Micro Tool Of Size Φ5 µm & 1mm Length	Tool size influence the accuracy
6	Yang et al & 2007	Semi Cylindrical Tool With Ultrasonic Vibrations	Improves the flow space of the electrolyte and electrolyte
7	Tsui et al & 2008	Helical Tool	Improves micro drilling accuracy
8	Wang et al & 2010	Wedged Shaped Electrodes	High quality inclination
9	Fan et al & 2010	Rotational Tool Electrode	Good for deep micro hole drilling
10	Hewidy et al & 2001	Orbital Tool Motion	Improves the accuracy of the hole

Sources: Authors Compilation

Although many researchers have already initiated some research in this area still in-depth research is required to improve the material removal rate and accuracy. Considering these requirements an EMM set-up is indigenously developed influence of tool electrode design (flat, conical with rounded & wedged electrode) and electrolyte concentration on machining rate and overcut were studied. Based on the previous experiments conducted the machining voltage, pulse on-time, frequency and machining current was standardized as 8V, 7.5ms, 50Hz and 0.8A.

EXPERIMENTAL SET-UP

The developed EMM set-up is shown in Fig. 1. The set-up basically consists of various elements and sub-systems. The mechanical machine unit consists of main machining body, tool electrode feeding attachment, work holding fixture and machining chamber.

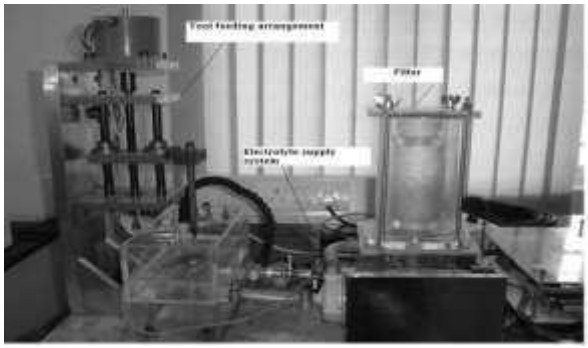
Main machine body is made up of mild steel and is chromium plated for corrosion resistance. The tool electrode feed mechanism, with resolution of 2 µm, along Z axis are designed with stepper motor and 8051 microcontroller.

The set-up maintains the initial set gap constantly throughout the machining by providing appropriate command signals to the stepper motor with the help of microcontroller unit. The current sensor continuously monitors the gap current in the closed loop and provides feedback to the microcontroller unit.

The work holding fixture and machining chamber is made up of electrically non-conductive perspex material. The electrolyte supply and cleaning system consist of a pump and filter.

A pulsed power supply of 20 V and 30 A with the capability for varying voltage, current and pulse width to the required values within the specified range, was used.

Figure-1: EMM Setup



Sources: Authors Compilation

The experiments were conducted with flat, conical with rounded and wedged stainless steel electrode of Φ 464 µm as shown in Fig. 2 a-c. The Olympus optical microscope is used for taking the images. The figure 2 a-c presented here is of 10X magnification. Sodium nitrate of varying concentrations are used as electrolyte, it is preferred for low throwing power. These are non-corrosive and have good machining accuracies, but with low electrical current efficiencies. The side walls of the tool electrodes are coated with a bonding liquid. 304 stainless steel of thickness 200 µm was used as the workpiece. The tool electrode is made as a cathode and workpiece is made as anode.

Figure-2: (a) Flat (b) Conical with Rounded (c) Wedged Electrode

In the EMM process the preparation of brine solution plays an important role. Brine solution was prepared by adding required grams of sodium nitrate with distilled water in a beaker. Then it is mixed properly with the help of a magnetic stirrer. The electrolyte is allowed to flow with low pressure so that fresh electrolyte is circulated by the electrolyte supply system into the machining chamber. Table3 shows the scheme of machining parameters used for the experiments.

Table-3: Scheme of Machining Parameters Used for Experiments

S.No.	Tool electrode design	Machining voltage(V)	Pulse on- time (ms)	Electrolyte concentration (mole/l)	Machining rate (µm/sec)	Overcut (µm)
1	Conical with rounded	8	7.5	20	0.059	11.25
2	Conical with rounded	8	7.5	25	0.077	28.35
3	Conical with rounded	8	7.5	30	0.152	38.55
4	Flat electrode	8	7.5	20	0.040	66.25
5	Flat electrode	8	7.5	25	0.054	78.18
6	Flat electrode	8	7.5	30	0.132	85.36
7	Wedged electrode	8	7.5	20	0.070	-
8	Wedged electrode	8	7.5	25	0.111	-
9	Wedged electrode	8	7.5	30	0.167	-

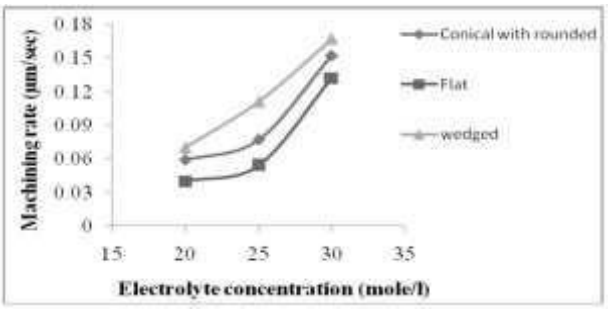
Sources: Authors Compilation.

EXPERIMENTAL ANALYSIS

Effect of Electrolyte Concentration on Machining Rate

Figure3 shows the effect of electrolyte concentration on machining rate. Based on the Figure3 the machining rate increases with increase in electrolyte concentration. Due to increase in electrolyte concentration the number of ions associated in the machining process increases. This increase in number of ions increases the current density between the tool electrode and the workpiece. According to Faraday's law the material removal rate is proportional to the current density. Hence, the machining rate increases with increase in electrolyte concentration.

Figure-3: Variation of Machining Rate with Electrolyte Concentration (Machining Voltage 8v, Pulse On-Time 7.5 Ms, Frequency 50Hz)



Sources: Authors Compilation.

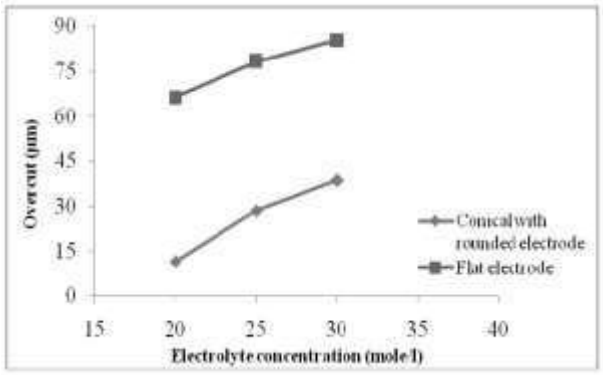
From the Fig. 3 the machining rate is lesser in flat electrode when compared with the wedged electrode. When flat electrode is used the flow of electrolyte to the machining zone is insufficient and hence the required amount of electrolyte is not supplied to the machining area which results in lower machining rate. The inadequate supply of electrolyte solution could not remove the products of the reaction and heat produced from machining region. Moreover, the non-localization of the current density distribution toward the workpiece also reduces the machining rate. The machining rate gradually increases in the range of 20-25 mole/l and rapidly increases in the range of 25-30 mole/l. For wedged electrode the machining rate is higher when compared with other two tool electrode with lesser accuracy. The nature of the curve is almost linear because the increment rate of dissolution efficiency is almost constant. The current density distribution in the case of the wedged electrode is narrow which improves the localization of the current resulting in higher machining rate. In case of conical with rounded electrode the shape provides the proper supply of the electrolyte solution to the machining zone compared to the flat electrode. The trend of the curve is similar to the flat electrode, between 25-30 mole/l the increase in dissolution rate reaches maximum.

Effect of Electrolyte Concentration On Overcut

Fig4 shows variation of overcut with electrolyte concentration. The overcut increases with increase in electrolyte concentration. When the concentration of electrolyte solution increases the ions associated with the machining process increases. This increase in ions improves the current density and resulting in more material removal. The reaction product resulting from machining has to be flushed out properly. The improper flushing of the dissolution products generates micro

sparks and removes the extra material from the workface resulting in higher overcut. Moreover the non-localization of the current in the machining zone results in higher overcut.

Figure-4: Variation of Overcut with Electrolyte Concentration (Machining Voltage 8v, Pulse On-Time 7.5 ms, Frequency 50Hz)



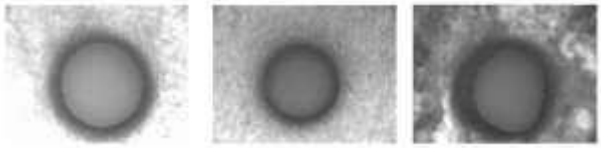
Sources: Authors Compilation.

Based on the Fig. 4 the overcut for flat electrode is more when compared with conical with rounded electrode. Due to the small gap between the front face of the tool and the workpiece surface, the current concentration inside the gap is much higher than in neighborhood. The dissolved materials are not removed due to the insufficient flow of electrolyte towards the inner region of the machining area. In the case of the conical with rounded electrode the front face area of the tool electrode is small and hence the profile of the electrode aids the flushing of the dissolved materials thus reduces the overcut. Fig. 5 a-c shows the optical microscope image at 10 x magnifications. It is evident from the Fig. 5 a-b that the overcut is more in the flat electrode when compared with conical electrode. The use of conical with rounded electrode improves the micro hole accuracy is by 2.2 times compared with flat electrode at electrolyte concentration of 30 mole/l, machining voltage 8v, pulse on-time 7.5 ms, frequency 50Hz.

Fig. 5c shows the optical image of the microhole machined using wedged electrode. In wedged electrode, the electrolyte solution distribution along the wedged side will be more compared to other side. It is evident from the Fig.4 that, the non-uniformity of electrolyte flow towards the machining zone reduces the accuracy and resulting in elliptical hole. So, it can be concluded that the wedge side supplies sufficient amount of electrolyte solution and others.

Figure: 5 shows the optical microscope image of micro hole at 10x magnification at machining voltage 8v, pulse on-time 7.5 ms, frequency 50Hz, electrolyte concentration of 30 mole/l, (a flat electrode, b. conical with rounded electrode, c. wedged electrode)

(a) (b) (c)



Sources: Authors Compilation.

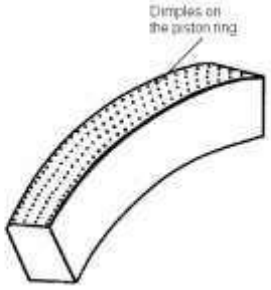
APPLICATIONS OF MICRO-MACHINING

The micromachining techniques find applications in various industries such as communications, semi-conductors, medicine, and ultra-precision machinery. Processing of advance microelectronic components, including Cu chips, high end packages and interconnects, thin film magnetic heads micro electro mechanical systems ,surface finish of print bands, Nozzle plates for ink jet printers, grooves for self-acting fluid film bearings, monolithic accelerometer and metal masks5. The novel application of electrochemical micromachining technology in automobile is as follows:

Automobile Applications

In diesel engines the engine performance can be improved by increasing the compression ratio. But this has the limited possiblilties at present because of friction,sealing and wear issues at the ring / cylinder linear interface.In deisel engine upto 60% of 15% mechanical losses resulting from friction between the piston rings and cylinder walls25. It has been estimated that reducing the frictional loss by even 10% can lead to a decreses in fuel consumption of upto 3%. The use of piston rings with dimples on the outer circumference improves the performance of the engine. The dimple on the piston rings contains the lubricant consequently the friction is reduced considerably. Electrochemical micromachining techniques can be applied to produce micro dimples on the piston rings as shown in the Fig 6. The proper dimensions of dimples, the friction performance of piston rings can be improved, particularly for light oils. The secondary lubrication effect in the dimpled area is the main mechanism responsible for improvement of performance.

Figure-6: Dimples on Piston Rings



Sources: Authors Compilation.

CONCLUSIONS

- The machining rate and overcut increases with increase in electrolyte concentration when using flat, conical with rounded and wedged electrode.
- Conical electrode produces 2.2 times lesser overcut and wedged electrode produces 1.26 times higher overcut, when compared with flat electrode at electrolyte concentration 30 mole/l, machining voltage 8v, pulse on-time 7.5 ms, frequency 50Hz.
- Wedged electrode produces elliptical hole for same machining conditions compared with flat and conical electrode. The wedged side supplies sufficient amount of electrolyte solution and other side it restricts the flow, resulting in improper machining.
- The application of EMM in piston rings improves the performance of the diesel engine by reducing the friction.

REFERENCES

[1] Yi, Qin; Brockett, A.; Ma, Y.; Razali, A.; Zhao J.; Harrison, C.; Pan, W.; Dai, X., and Loziak, D., (2010). Micro-Manufacturing: Research; Technology Outcomes and Development Issues", International Journal of Advertising Mfg Tech, Volume: 47, pp. 821–837.

[2] Bhattacharyya, B.; Munda, J.; and Malapati, M., (2004). "Advancement in Electrochemical Micromachining", Journal Mach Tools Manuf', Volume: 44, pp. 1577-1589.

[3] Jain, V. K., (2002). Advanced Machining Processes, New Delhi: Allied Publishers Pvt. Ltd.

[4] Datta, M., and Landolt, D., (2000). "Fundamental Aspects and Applications of Electrochemical Micro-Fabrication, 'Electrochemica Acta', Volume: 45, pp. 2535-2558.

[5] Bhattacharyya, B.; Mitra, S.; and Boro, A. K., (2002). "Electrochemical Machining: New Possibilities for Micromachining", J Robo and Comp Int Manuf, Volume: 18, pp. 283 - 289.

[6] Yong, L.; Yunfei, Z.; Guang, Y.; and Liangqiang, P., (2003). "Localized Electrochemical Micromachining with Gap Control", Sens Actuators A, Volume: 108, pp. 144-148.

[7] Kozak, J.; Rajurkar, K. P., and Makkar, Y., (2004). "Selected Problems of Micro Electrochemical Machining", Journal of Material Processing Technology, Volume: 149, pp. 426-431.

[8] Kurita, T.; Chikamori, K.; Kubota, S., and Hattori, M., (2006). "A Study of Three-Dimensional Shape Machining With An ECμM System", J Mach Tools Manuf, Volume: 46, pp. 1311-1318.

[9] Zhao, W.; Li, X., and Wang, Z., (2006). "Study on Micro Electrochemical Machining at Micro to Meso-Scale", In: Proceedings of the 1st IEE International Conference on Nano/Micro Engineered and Molecular Systems, Zhuhai, China, January 18th -21st, pp. 325-329.

[10] Lipiec, P.; Skoczypiecl, S.; Wyszynski, D., and Ruszaj, A., (2007). „Primary Experimental Research Of Electrochemical Micromachining Process", In: Proceedings of 15th International Symposium on Electromachining, Pittsburg, Pennsylvania, USA, April 23rd-27th, pp. 335-358.

[11] Zhang, Z., and Zhu, D., (2008). „Experimental Research on the Localized Electrochemical Micro-machining", Russian Journal of Electrochemistry', Volume: 44, pp. 926-930.

[12] Tsui, H. P.; Hung, J. C.; You, J. C., and Yan, B. H., (2008). „Improvement of Electrochemical Microdrilling Accuracy using Helical Tool", J Mater and Manuf Process, Volume: 23, pp. 499-505.

[13] Ryu, S. H., (2009). "Microfabrication by Electrochemical Process in Citric Acid Electrolyte", J Mater Process Technol, Volume: 209, pp. 2831-2837.

[14] Ahn, S. H.; Ryu, S. H.; Choi, D. K., and Chu, C. N., (2004). „Electro-chemical Micro Drilling Using Ultra Short Pulses", Precision Engg', Volume: 28, pp. 129-134.

[15] Westley, J. A.; Atkinson, J., and Duffield, A., (2004). „Generic Aspects of Tool Design for Electrochemical Machining", 'J Mater Process Technol', Volume; 149, pp. 384-392.

[16] Ebeid, S. J.; Hewidy, M. S.; El-Taweel, T. A., and Youssef, A. H., (2004). "Towards Higher Accuracy for ECM Hybridized with Low-Frequency Vibrations Using the Response Surface Methodology", J Mater Process Technol, Volume: 149, pp. 432-438.

[17] Lee, E. S.; Baek, S. Y., and Cho, C. R., (2007). „A Study of the Characteristics For Electrochemical Micromachining With Ultrashort Voltage Pulses", Intl Journal of Adv Mfg Tech, Volume: 31, pp. 762–769.

[18] Yang, I.; Park, M .S., and Chu, C. N., (2007). "Micro ECM With Ultrasonic Vibrations Using Semi-Cylindrical Tool", Int J. Precis Engg Manuf', Volume: 10, pp 5-10.

[19] Wang, W.; Zhu, D.; Qu, N.; Huang, S.; and Fang, X., (2010). "Electrochemical Drilling Inclined Holes Using Wedged Electrodes", Intl Journal of Adv Mfg Tech, Volume: 47, pp. 1129–1136.

[20] Fan, Z. W., and Hourng, L. W., (2010). "Electrochemical Micro-Drilling of Deep Holes by Rotational Cathode Tools", Intl Journal of Adv Mfg Tech, Volume: 52, pp. 555-563.

[21] Hewidy, M. S.; Ebeid, S. J.; Rajurkar, K. P., and El-Safti, M. F., (20010. "Electrochemical Machining under Orbital Motion Conditions", 'J Mater Process Technol, Volume: 109, pp. 339- 346.

[22] Xiaobin, Lu, and M., M. Khonsari, "An Experimental Investigation of Dimple Effect on the Stribeck Curve of Journal Bearings", Tribology Letters, Volume: 27, Number: 2, pp. 169-176, DOI: 10.1007/s11249-007-9217-x.

*****

QUANTUM CHEMICAL STUDY OF SOME CYCLIC NITROGEN COMPOUNDS AS CORROSION INHIBITORS

P.Vidhya[127] V. Kannapan[128] V. Sathyanarayanamoorthy[129]

ABSTRACT

The molecular structure of Pyrazole, 1 Nitropyrazole have been investigated by using abinitio Hartree Fock and density Functional Theory(DFT) using standard B3LYP functional and 6-311G(d)and6-311G(d,p) basis sets. Corrosion control can be achieved by many methods, being corrosion inhibitors one of the most effective alternatives for the protection of metallic surfaces against corrosion. A perusal of the literature on corrosion inhibitors reveals that most organic inhibitors employed as corrosion inhibitors contain nitrogen, oxygen, sulphur and/or aromatic ring in their molecular structure.

The aim of this paper is to extend these investigations in order to discuss the relationship between quantum chemical calculations and experimental inhibition efficiencies of the inhibitors by determining the quantum chemical parameters such as the energies of highest occupied molecular orbital (EHOMO) and the lowest unoccupied molecular orbital (ELUMO), the energy difference (ΔE) between EHOMO - and ELUMO and dipole moment (μ). The chemical structures of the compounds studied are given in Figure.

The optimized molecular structures of the studied molecules using hybrid DFT functional (B3LYP/6-31G*) and the calculated quantum chemical indices EHOMO, ELUMO ΔE and dipole moment (μ) are given in table1,2, 3. The atomic charge values were obtained by the Mulliken population analysis. Table 4 & 5 presents Mulliken charges of the selected atoms of the compounds studied.

KEYWORDS

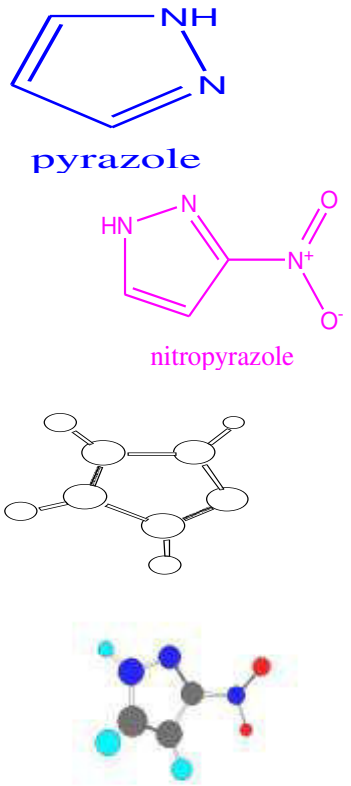
Pyrazole Derivatives, Quantum Chemical Parameters, and Mullikan Atomic Charges etc.

INTRODUCTION

The term Pyrazole was given by Ludwig Knorr in 1883. Pyrazole (Fig.1) refers to the class of simple aromatic ring organic compounds of the heterocyclic series characterized by a 5-membered ring structure composed of three carbon atoms and two nitrogen atoms in adjacent positions. Being so composed and having pharmacological effects on humans, they are classified as alkaloids, although they are rare in nature. In 1959, the first natural pyrazole, 1-pyrazolyl-alanine, was isolated from seeds of watermelons [1].

Pyrazole derivatives have a long history of application in agrochemicals and pharmaceutical industry as herbicides and active pharmaceuticals. The recent success of pyrazole COX-2 inhibitor has further highlighted the importance of these heterocyclic rings in medicinal chemistry. A systematic investigation of this class of heterocyclic lead revealed that pyrazole containing pharmacoactive agents play important role in medicinal chemistry. The prevalence of pyrazole cores in biologically active molecules has stimulated the need for elegant and efficient ways to make these heterocyclic lead [1].

Figure-1: Structure of pyrazole and 1-Nitropyrazole & molecular structure of pyrazole and 1-Nitropyrazole



Sources: Authors Compilation

METHOD OF CALCULATION

The compound under investigation is purchased from sigma-aldrich chemicals, U.S.A. which is of spectroscopic grade and hence used for recording the spectra as such without any further purification.The FTIR spectra of the compounds are recorded in the range of 4000-100cm^{-1},with scanning speed of 30cm^{-1}min^{-1}. The frequencies of all sharp bands are accurate to ± 1 cm^{-1}

Quantum Chemical Calculations

Abinitio (HF/6-31G**basis set) calculations were done by the GAMESS program Suite (2). Bulk solvent effects were estimated by single point calculations using the polarized continuum model (PCM)(3,4).

[127] VSB College of Engineering Technical Campus, Tamil Nadu, India, vidhyapadmanabhan28@gmail.com
[128]Presiency College, Tamil Nadu, India, vkann@gmail.com
[129] PSG College of Arts and Science, Tamil Nadu, India, sathyanarayanamoorthi@yahoo.co.in

Quantum Chemical Parameters

Quantum chemical methods and molecular modeling techniques enable the definition of a large number of molecular quantities characterizing the reactivity, shape, and binding properties of a complete molecule as well as of molecular fragments and substituents. The use of theoretical parameters presents two main advantages: firstly, the compounds and their various fragments and substituents can be directly characterized on the basis of their molecular structure only; and secondly, the proposed mechanism of action can be directly accounted for in terms of the chemical reactivity of the compounds under study [5].

Molecular Orbital Energies

Highest occupied molecular orbital energy (E_{HOMO}) and lowest unoccupied molecular orbital energy (E_{LUMO}) are very popular quantum chemical parameters. These orbitals, also called the frontier orbitals, determine the way the molecule interacts with other species. The HOMO is the orbital that could act as an electron donor, since it is the outermost (highest energy) orbital containing electrons. The LUMO is the orbital that could act as the electron acceptor, since it is the innermost (lowest energy) orbital that has room to accept electrons. According to the frontier molecular orbital theory, the formation of a transition state is due to an interaction between the frontier orbitals (HOMO and LUMO) of reactants [7].

Dipole Moment

The most widely used quantity to describe the polarity is the dipole moment of the molecule [8]. Dipole moment is the measure of polarity of a polar covalent bond. It is defined as the product of charge on the atoms and the distance between the two bonded atoms.

Semiempirical Methods

Semiempirical methods differ in the details of the approximations (e.g. the core-core repulsion functions) and in particular in the values of the parameters. The semiempirical methods can be optimized for different purposes. The MNDO, AM1 and PM3 methods were designed to reproduce heats of formation and structures of a large number of organic molecules. Other semiempirical methods are specifically optimized for spectroscopic properties (e.g. INDO/S or CNDO/S).

MNDO (Modified Neglect Of Differential Overlap)

It is based on the NDDO (neglect of diatomic differential overlap) approximation and in turn NDDO is an improved version of INDO (intermediate neglect of differential overlap) method. INDO itself is an improvement over the CNDO (complete neglect of differential overlap) approximation. There are several such semiempirical LCAO MO methods, developed for specific purposes.

AM1 (Austin model 1)

AM1 is a semiempirical method based on the neglect of differential diatomic overlap integral approximation. Specifically, it is a generalization of the modified neglect of diatomic differential overlap approximation. AM1 was developed by Michael Dewar and coworkers reported in 1985 [10].

PM3 (Parameterized model number 3)

PM3 is another semiempirical method based on the neglect of differential diatomic overlap integral approximation. The PM3 method uses the same formalism and equations as the AM1 method.

Corrosion Inhibitors Studied by Semiempirical Methods

The inhibition of corrosion in acid solutions can be affected by the addition of a variety of organic molecules. Compounds containing nitrogen, oxygen and sulphur have shown vast applications as corrosion inhibitors. The studies of Vosta and Eliasek [11] and Chakrabarti can be respected as the first theoretical studies on corrosion inhibitors in the literature. Subsequently, Costa and Lluch reported the results of quantum mechanical calculations of different compounds, such as linear-chain diols, diamines and aliphatic aminoalcohols as corrosion inhibitors of metals. They calculated HOMO and LUMO energies by using the semiempirical MINDO/3 method. A good agreement was found between their theoretical predictions and experimental data. Sayos et al. performed theoretical calculations on several aromatic organic compounds as corrosion inhibitors for titanium and its alloys employing the MNDO method. They also concluded that the agreement with the experimental data was quite good. In the following decades, several studies on various compounds were reported which dealt with correlations of quantum chemical properties obtained with semiempirical tools with experimental findings. For instance, the influence of some heterocyclic compounds, i.e. some oxadiazole derivatives, on the corrosion of mild steel in acid solutions, has been investigated by Lagrenee et al, and Bentiss et al. Beside using experimental methods, they used AM1 semiempirical method to obtain the electronic properties of those compounds. They calculated E_{HOMO}, E_{LUMO}, $\Delta E(E_{LUMO} - E_{HOMO})$ and dipole moment (μ) and found a highly significant multiple correlation coefficient between experimental and theoretical data.

Semiempirical calculations to evaluate the efficiency of some imidazole derivatives as acidic corrosion inhibitors for zinc and iron have been performed by Bereket et al. [12] using AM1, PM3, MNDO and MINDO/3 methods. Charges on nitrogen atoms, total energy, ionization potential E_{HOMO}, E_{LUMO}, ΔE (E_{LUMO}- E_{HOMO}) and dipole moment (μ) have been calculated and correlated with experimental results. A satisfactory agreement was found between theoretical and experimental data. Similar studies on these compounds have been carried out by ogretir et al. in order to probe the inhibition mechanism of corrosion via metal-ligand interaction using semiempirical methods. According to their conclusion, semiempirical calculations can be used to elucidate the mechanism of inhibition.

RESULT AND DISCUSSION

Corrosion control can be achieved by many methods, being corrosion inhibitors one of the most effective alternatives for the protection of metallic surfaces against corrosion. A perusal of the literature on corrosion inhibitors reveals that most organic inhibitors employed as corrosion inhibitors contain nitrogen, oxygen, sulphur and/or aromatic ring in their molecular structure. The aim of this paper is to extend these

investigations in order to discuss the relationship between quantum chemical calculations and experimental inhibition efficiencies of the inhibitors by determining the quantum chemical parameters such as the energies of highest occupied molecular orbital (E_{HOMO}) and the lowest unoccupied molecular orbital (E_{LUMO}), the energy difference (ΔE) between E_{HOMO} and E_{LUMO} and dipole moment (μ).

In order to find optimized conformations of the compounds studied and to speed up the calculations, the molecular structures were optimized initially with PM3 semiempirical calculation. The convergence was set to 0.001 $Kcalmol^{-1.}$ The structures obtained from PM3 calculation were fully reoptimized by using DFT (density functional theory) methods to estimate the quantum chemical parameters. Calculations at the DFT level were performed using one basis sets, 6-31G(d).

The chemical structures of the compounds studied are given in Figure. The optimized molecular structures of the studied molecules using hybrid DFT functional (B3LYP/6-31G*) and the calculated quantum chemical indices $E_{HOMO,}$ E_{LUMO} ΔE and dipole moment (μ) are given in table 1,2,3.

According to the frontier molecular orbital theory, the formation of a transition state is due to an interaction between frontier orbitals (HOMO and LUMO) of reacting species. Thus, the treatment of the frontier molecular orbitals separately from the other orbitals is based on the general principles governing the nature of chemical reactions. HOMO is often associated with the electron donating ability of a molecule. High E_{HOMO} values indicate that the molecule has a tendency to donate electrons to appropriate acceptor molecules with low energy empty molecular orbital. Increasing values of the E_{HOMO} facilitate adsorption (and therefore inhibition) by influencing the transport through the adsorbed layer. E_{LUMO} indicates the ability of the molecules to accept electrons. The lower values of E_{LUMO}, the more probable it is that the molecule would accept electrons. Low absolute values of the energy band gap (ΔE) gives inhibition efficiencies, because the energy to remove an electron from the last occupied orbital will be low. The results obtained by 6-311G(d,p)/B3LYP/6-311G(d,p) method (Table 1) show that nitropyrazole has the highest HOMO energy (E_{HOMO}) and the lowest LUMO energy (E_{LUMO}) among these organic heterocyclic compounds. Another point to be considered is the HOMO-LUMO gap (ΔE) i.e, the difference between the HOMO and LUMO energies for the compounds. Nitro pyrazole has the smaller energy gap as compared to other molecules. According to these theoretical results and the experimentally found inhibition efficiencies, it can be said that nitro pyrazole has more inclination to get adsorbed on the metal surface than PYRAZOLE. The binding capability of a molecule with metal depends also on the electronic charge on the chelating atom. Thus, the atomic charge values were obtained by the Mulliken population analysis. Table 4 & 5 presents Mulliken charges of the selected atoms of the compounds studied. From the atomic charge values listed, 1C atom of ***nitro pyrazole*** has excess electron density which increases the π- electron density in the aromatic ring. It is well known that the chelating or binding capability of a molecule with a metal depended on the electronic charge on the chelating or active atoms, i.e. the more negative the charge, the stronger of the binding capability. The dipole moment (μ) is another indicator of the electronic distribution in a molecule and is one of the properties used to discuss and to rationalize the structure. No significant relationship has been found between the dipole moment values and inhibition efficiencies. Besides, there is a lack of agreement in the literature on the correlation between the dipole moment and inhibition efficiency.

Calculated Quantum Chemical Parameters of Studied Compounds

Table-1: AM1

Molecule	E(HOMO) Kcal/mol	E(LUMO) Kcal/mol	ΔE Kcal/mol	ΔH Kcal/mol	μ Debye
Pyrazole	-1.52	0.036	-31.28	65.56	2.11
Nitropyrazole	-1.37	0.098	-54.22	49.33	11.41

Sources: Authors Compilation.

Table-2: PM3

Molecule	E(HOMO) Kcal/mol	E(LUMO) Kcal/mol	ΔE Kcal/mol	ΔH Kcal/mol	μ Debye
Pyrazole	-1.41	0.056	-27.21	48.83	2.37
Nitropyrazole	-1.26	0.11	-48.09	50.80	9.83

Sources: Authors Compilation.

Table-3: DFT

Molecule	E(HOMO) Kcal/mol	E(LUMO) Kcal/mol	ΔE Kcal/mol	ΔH Kcal/mol	μ Debye
Pyrazole	-15.64	0.171	-224.70	-	2.64
Nitropyrazole	-18.93	0.0041	-375.54	-	9.047

Sources: Authors Compilation.

Table-4&5: Total Mulliken Atomic Charge Pyrazole &Total Mulliken Atomic Charge 1-Nitropyrazole

Atom	Mulliken Population	Charge
1C	7.68	-0.68
2C	5.83	0.18
3C	6.37	-0.37
4N	7.26	-0.26
5C	5.98	0.030
6H	0.60	0.41
7H	0.76	0.25
8H	0.79	0.22
9H	0.78	0.23

Sources: Authors Compilation

Atom	Mulliken Population	Charge
1C	6.07	-0.07
2C	6.20	-0.20
3N	7.39	-0.39
4C	5.97	0.039
5H	0.73	0.28
6H	0.89	0.12
7H	0.90	0.11
8N	6.89	0.12
9O	8.51	-0.51
10O	8.50	-0.50

Sources: Authors Compilation

CONCLUSIONS

Through PM3 semi empirical and DFT quantum chemical calculations a correlation between parameters related to the electronic structure of some cyclic nitrogen compounds and their ability to inhibit the corrosion process could be established. The highest occupied molecular orbital energy levels and energy gaps calculated by 6-311G (d,p) DFT study show reasonably good correlation as compared to other calculated data. Comparison of theoretical and experimental data exhibit good correlation confirming the reliability of the method employed here.

REFERENCES

[1] Eicher, T., and Hauptmann, S., (2003), Edition IInd, 'The Chemistry of Heterocycles: Structure, Reactions, Syntheses, and Applications', Wiley-VCH, ISBN 3527307206.

[2] Frisch, M. J., et al., Gaussian Inc.: Pittsburgh PA 1998.

[3] Miertus, S.; Scrocio. E., and Tomasi, J., (1981). Chem. Phys, 55:117

[4] Barone, V.; Cossi, M., and Tomasi, J., (1998). J. Comput. Chem, 19:404.

[5] I., N. Levine, (1991). Quantum Chemistry, New Jersey: Prentice Hall.

[6] K. Fukui, Theory of Orientation and Stereoselection, Springer-Verlag, NewYork, 1975.

[7] O., Kikuchi, (1987). "Systematic QSAR Procedures with Quantum Chemical Descriptors", Quant. Struct. Act. Relat. 6, pp. 179–184.

[8] P., Hohenberg, and W., Kohn, "Inhomogeneous Electron Gas", Phys. Rev., 136 (1964), B864–871.

[9] M., J. S. Dewar; E., G. Zoebisch; E., F. Healy, and J. J. P. Stewart, (1985). Development and Use of Quantum Mechanical Molecular Models.76.AM1: A New General Purpose Quantum Mechanical Molecular Model, J. Am. Chem. Soc. 107, pp. 3902–3909.

[10] G., Bereket, and C., Ögˇretir, and A., Yurt, (2001). "Quantum Mechanical Calculations on some 4-Methyl-5-Substituted Imidazole Derivatives as Acidic Corrosion Inhibitor for Zinc", J. Mol. Struct. (THEOCHEM) 571, pp. 139–145.

*****

DESIGN AND ANALYSIS OF HORIZONTAL AXIS WIND TURBINE BLADE

N. Prabakaran[130] Kannan Jegathala Krishnan[131]
S. Dhamodharan[132]

ABSTRACT

The wind turbine blade is a very important part of the rotor.Exrraction of energy from wind depends on the design of blade. In this work a blade of length 1.5m for 1kw horizontal axis wind turbine is designed by BEM (Blade Element Momentum) theory.

A computer programme is developed to get dimension (Twist, chord and thickness). The aerofoil taken for the blade is S809 & S804. The airfoil taken is root to tip. The analysis of designed blade is done in ANSYS.Result obtained from ANSYS is compared with the theoretical work.

KEYWORDS

Design, Chord, Twist, Blade etc.

INTRODUCTION

Wind turbines are subjected to very specific loads and stresses. Due to the nature of wind, loads are highly variable. Varying loads are more difficult to handle than Static loads because the material becomes fatigued. Moreover as a working medium the air is of low density so that the surface required for capturing energy must be large. When designing a wind turbine, the aim is to attain the highest possible power output under particular atmospheric conditions and this depends on the shape of the blade.

The change of the shape of blade is one of methods to modify stiffness and stability, but it may influence aerodynamic efficiency of wind turbine. Other method to change dynamic and mechanical properties of wind turbine is modifying the composite material, which the blade is made off.

Motivation

With the recent surge in fossil fuel prices, demand for cleaner energy sources, and government funding incentives, wind turbines have become viable technology for power generation. Currently horizontal axis wind turbine dominates the market, mainly because they have high aerodynamic efficiency when compared to vertical axis turbines. Small scale HAWT act as alterrnate source of power at times of power failures. Therefore, research and development of small scale HAWT is much important for wind power generation. It is decided to develop a model of horizontal axis wind turbine and to calculate its power co-efficient by means of aerodynamic calculation along with computational analysis.
Smallwindturbines

All wind turbines can be characterized as either Horizontal Axis Wind Turbines (HAWT) or Vertical Axis Wind Turbines (VAWT). In HAWT, the rotor spins about an axis horizontal to the earth's surface. The rotor of a VAWT spins about an axis perpendicular to the Earth's surface. Vertical axis wind turbines are typically developed only for built environment. Changes in wind direction have fewer negative effects on this type of turbine because it does not need to be positioned into the wind direction. However, the overall efficiency of these turbines in producing electricity is lower than HAWT. VAWTs are categorized as Savonius or Darrieus types, according to the principle used to capture the wind flow. For the Savonius type, the wind pushes the blades, which implies that the rotation speed is always lower than the wind speed. Contrary to that, the shape of the rotor of the Darrieus type makes it possible for the rotor to spin faster than the wind speed.

Rotors of HAWT are placed on towers to position them where the wind speed is fastest and exhibits most power (Jonkman, 2003). A nacelle typically resides atop the tower and contains the support structure for the rotor, the rotor shaft, a gearbox and the electric generator (alternator). The gearbox is used to transform the low-speed high-torque power of the rotor to high-speed, low-torque power that can run the electric generator.

Rotor can be positioned upwind or downwind of the tower. Downwind rotor configurations can track the wind automatically as wind direction changes. However, the wind must flow around the tower to reach the rotor of a downwind turbine. This results in complex flow patterns and periodic fluctuations in aerodynamic loads, which have important dynamic effects on the turbine structure. Flow passing through the rotor plane is unobstructed by the tower for upwind rotor configurations.

The hub structure connects the blades to the drive shaft. Hubs are generally characterized as either rigid or teetering. In rigid hub designs, the hub is rigidly attached to the drive shaft. In contrast teetered hubs are connected to the drive shaft by means of a teeter pin, a bearing that permits the rotor to rock into and out of the plane of rotation. Teetered hubs have the benefit that bending moments brought about by thrust forces acting on the blades are not transferred to the nacelle and tower structure. Consequently, the nacelle and tower structures of turbines with rigid hub must be designed more robustly than those with teetered hub.

Small wind turbine need to be reliable, affordable and almost maintenance free. To meet these criteria, optimal turbine performance is sometimes sacrificed for simplicity in design and operation (Andrew, 2005). Thus, rather than using the generator as a motor to start and accelerate the rotor when the wind is strong enough to begin producing power, small wind turbines rely solely on the torque produced by the wind acting on the blades.

Furthermore, small wind turbines are often located where the generated power is required, which is not necessarily where the wind resource is best. In low or unsteady wind conditions slow starting potentially reduces the total energy generated. Also, a

[130] Assistant Professor, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, prabakaraned@gmail.com
[131] PG Scholar, School of Engineering & Science, Victoria University, Australia, Kannan.Jegathala-Krishnan@vu.edu.au
[132] Lecturer, Dhanalakshmi Srinivasan College of Engineering, Tamil Nadu, India, ersdhamotharan@gmail.com

stationary wind turbine fuels the perception of wind energy as an unreliable energy source.

The main technical challenge in the design of a small wind turbine is to come up with a system configuration and control algorithm that maximizes wind energy production from the turbine and also provide favourable charging conditions for batteries. This task is complex because of the variability of the wind, which results in varying wind turbine power output. Ideally, the system configuration and its control should optimize the match between the wind turbine rotor and load, thereby allowing the maximum available power from the wind to be used, while at the same time charging the batteries with an optimum charge profile (Corbus et al., 1999).

The generators of small turbines often cause a significant resistive torque that must be overcome aerodynamically before the blades will start turning. Furthermore, pitch control is rarely used on small wind turbine because of cost. Thus, it is not possible to adjust the turbine blade's angle of attack to the prevailing wind conditions. This problem is particularly acute during starting.

A further major difference is that small turbines usually operate with varying shaft speed in an attempt to maintain maximum performance as the wind speed varies. Many large turbines run at constant speed as this allows the generator to maintain synchronicity with the utility grid.

The IEC 61400-2 (International Electrotechnical Commission) defines a small turbine as having a swept area less than 200m^2, which correspond to a power output of about 120kW. In addition, there is a further subdivision in that turbine of swept area less than 2m^2 (about 1.2kW) do not need to have their tower included in the certification process (Introduction to Wind Turbine Technology:
accessed at http://www.wind.newcastle.edu.au/notes.html on 12th February, 2007).

Wind Turbine Blades

All forms of wind yurbine are designed to extract power from a moving air stream. The blades have an airfoil cross-section and extract wind dy lift force caused by a pressure difference between blade sides. When air passes over an airfoil section: it travels faster over the top of the top of the blade than it does below it. This makes the air pressure above the blade lower than it is below. To the unequal pressures the blade experiences a lifting force (BWEA). For maximum efficiency, the blades often incorporate twist and taper.

The mechanical power produced by a rotor is purely a function of the blade geometry and the incident velocity. The design parameters that affect aerodynamic performance include blade pitch (angle of attack), taper, and twist distribution. For a given blade, its geometric shape is usually fixed, i.e. the aerodynamic shape, taper and twist distribution do not change.

The torque produced by the rotor can be controlled in two ways: changing the geometry by varying the blade pitch angle, or by changing the rotor's rotational speed so that the rotor operates at the optimal blade tip speed ratio.

In the beginning most wind turbine blades were adaptations of airfoil developed for aircraft and were optimized for wind turbines uses. In recent years development of improved airfoil sections for wind turbines has been on going. The prevailing tendency is to use NACA 63(National Advisory Council on Aeronautics) and NREL S809 (National Renewable Energy Laboratory) airfoil cross-section that may have modifications in order to improve performance for special applications and wind conditions.

Design Specification

Chart-1

• Power	: 1000 watts
• Diameter	: 3 meter
• Tip speed ratio	: 5
• No of blades	: 3
• Rated wind speed	: 6 m/s
• Power co-efficient	: 0.45
• Profile used	: S 823 & S822

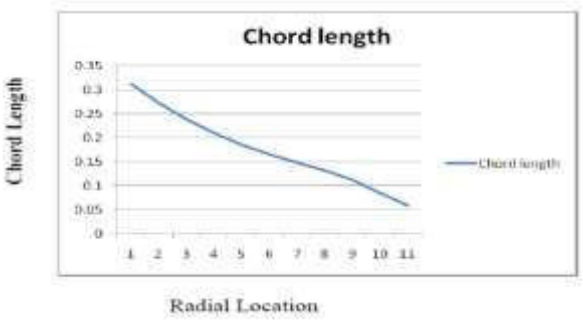
Sources: Authors Compilation

TWIST, CHORD AND THICKNESS DISTRIBUTION

The twist of a wind turbine blade is defined in terms of the chord line. It is a synonym for the pitch angle. However the twist defines the pitch settings at each station along the blade according to the local flow conditions. The pitch angle(β) is larger near the root (where local speeds are low),and small at the tip (where local speeds are high).The apparent wind angle changes along the blade due to the increase in blade speed with increasing distance outboard. Hence to maintain optimum angle of attack of the blade section to the wind, it must be twisted along its length. The twist distribution is maintained such that the lift coefficient will be maximum at every section. Chord direction is perpendicular to the span direction and lies in the plane extending through the laeding edge and the trailing edge.A shoulder is the point where chord is maximum and it is minimum at the tip of the blade. Thickness distribution is calculated in terms of the chord where the total thickness of the blade at any station will be percentage of the chord length at that station.

Figure1 shows the chord distribution for the blade. Figure2 shows the twist distribution for the blade. Both chord and thickness are reducing from root to tip.

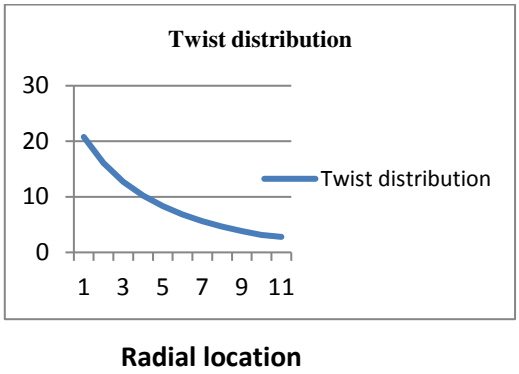
Graph-1: Radial Location VS Chord Length



Sources: Authors Compilation

Fig 3to7 indicates the Liftand Drag coefficient of two different aerofiles with different angle of attack.

Figure-2: Radial Location V/s Twist Distribution



Sources: Authors Compilation

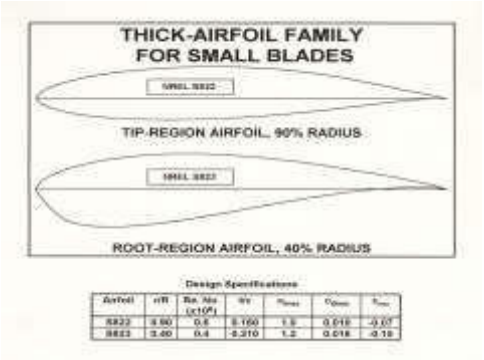
AIRFOIL SELECTION

In this thesis what has been considered in constructing the airfoil database from the computer program XFOIL which includes these considerations; two-dimensional data with Reynolds number in the order of 8 million. There are evidently many engineering requirements into the selection of a wind turbine airfoil. These include primary requirements related to aerodynamic performance, structural, strength and stiffness, manufacturability and maintainability. The usual assumption is that high lift and low drag are desirable for an airfoil and that the drag-to-lift ratio γ which is known as glide ratio as given below is a critical consideration

$$\gamma = \frac{C_D}{C_L}$$

Airfoils for HAWTs are often designed to be used at low angles of attack, where lift coefficients are fairly high and drag coefficients are fairly low(i.e. fairly low glide ratio). Many different standard airfoils developed for aircraft have been used on HAWTs. The NACA 230XX series and NACA44XX series airfoils have been used on many modern HAWT units. The NACA 63-2XX series airfoils have demonstrated the best overall performance characteristics of the NACA airfoil families.

Figure-3: Profile Drawing

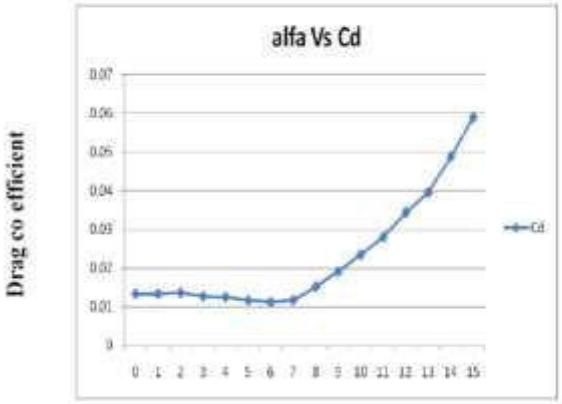


Sources: Authors Compilation

Angle of attack

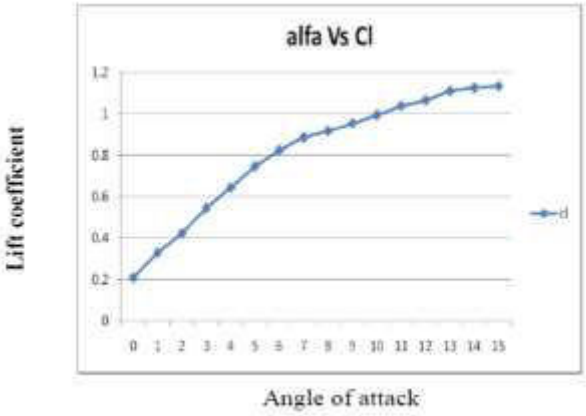
Fig.4. Angle of attack Vs Drag co efficient

a)S822 profile



Sources: Authors Compilation

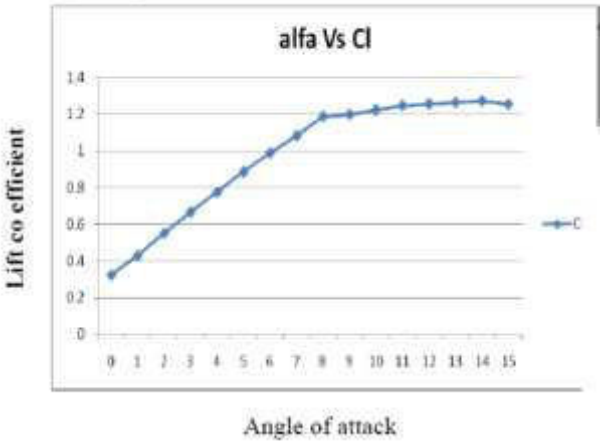
Fig .5. Angle of attack Vs Lift co efficient



Sources: Authors Compilation

Fig.6.Angle of attack Vs Lift co efficient

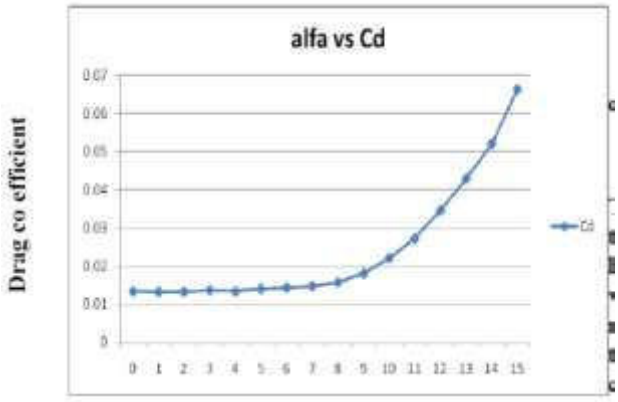
b) S823 profile



Sources: Authors Compilation

Angle of attack

Fig.7. Angle of attack Vs Drag co efficient



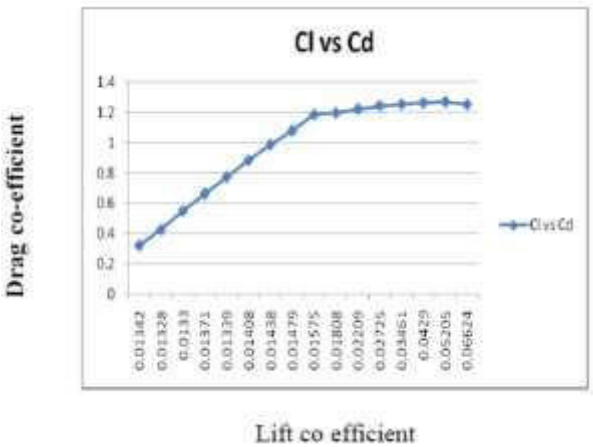
Sources: Authors Compilation

Fig.8. Lift co-efficient VS Drag co-efficient

c) SERI 822

Sources: Authors Compilation

Fig.9. Lift co-efficient VS Drag co-efficient



Sources: Authors Compilation

Figure 8 and 9 indicates the lift and darg co-efficient of 2 different profiles.

COMPUTATIONAL DESIGN

The horizontal axis wind turbine consist of one, two or three horizontal airfoil shaped blades connected with the hub, which rotates the horizontal shaft. When the wind velocity is high, large amount of aerodynamic forces and moments would act on the blade. Also high torque and twisting moments would exist on the shaft. This would create some mechanical failure and loss of efficiency. Therefore the design of hub and supports are also much important for higher efficiency and structural rigidity. The design aspects and parameters of every part are explained below.

A) Blade Design

The blade design is the root factor for the power generation. Therefore the selection procedure and design involves large amount of survey and calculation. With the help of aerodynamic calculation and performance, the airfoil S 822 and S 823 are selected for blade design. The profile of the airfoil is plotted using the data points. The complete design is done using CATIA V17.

Steps involved in blade design are listed below:

- Co-ordinates are calculated for S822 and S823 profiles for their respective chord lengths.
- Profiles are drawn at their respective radial location.
- Profiles are given angle of twist.
- Multi-session of these profiles is done to get single blade geometry.

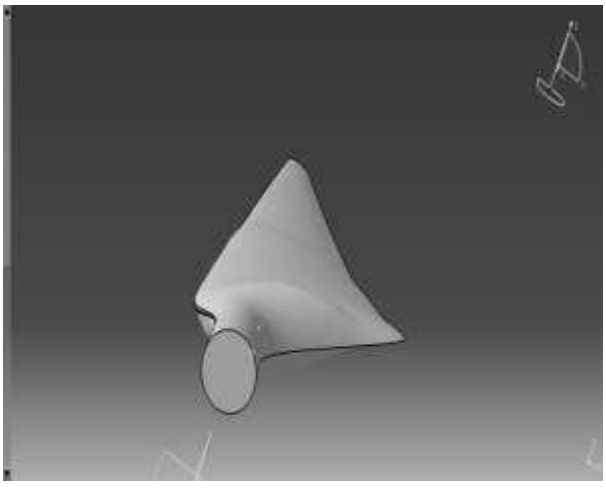
Rotor Specification:

Table-1

Airfoil	S 823(root) & S822(tip)
Blade length	1.5 m

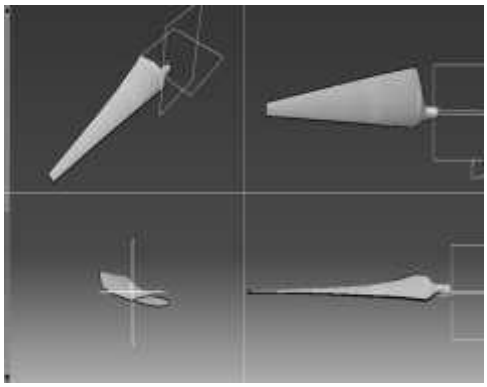
Sources: Authors Compilation

Figure-10: Designed Blade



Sources: Authors Compilation

Figure-11: Four Views of Designed Blade



Sources: Authors Compilation

STRUCTURAL ANALYSIS

The blade should have sufficient structural rigidity to withstand high stress developed at the dynamic condition. For that the finite element analysis is performed on the blade and structural efficiency is calculated. The 3-d model designed in CATIA is imported to ANSYS 11- the finite element analysis software.

The steps carried out in structural analysis are listed below:

- Importing the geometry from CATIA in .stp format.
- Specifying material type and applying material to the geometry.
- Meshing the geometry.
- Giving boundary conditions and applying the loads.
- Solving the problem and plotting the results

A) Forces Acting On the Blade:

The forces to be considered for the blade analysis are:

- The dynamic pressure force acting all over the length of the blade.
- The aerodynamic loads of lift, drag and its resultant forces.
- The centrifugal force due to rotation of the rotor at the dynamic condition.
- The self-weight of the blade.

B) Boundary Conditions and Loads:

The global coordinate system of the blade is designed in such a way that the airfoil shape lies in the y-z plane and the x-axis is along the length of the blade.

The following boundary conditions are applied for the simple steady state analysis of the blade.

- Since the root section of the blade is to be attached with the hub all the degrees of freedom are made zero at the root.
- The dynamic pressure, aerodynamic loads are applied uniformly all over the nodes.
- The centrifugal force is applied at the centre of gravity of the geometry

RESULTS AND DISCUSSIONS

Finite element analysis is performed on the rotor blades. The material, geometrical properties are given in the tables.

Table-2: Material Properties

Material	Aluminum alloy
Young's Modulus	7.1e+010 Pa
Poisson's Ratio	0.33
Density	2770. kg/m³
Thermal Expansion	2.3e-005 1/°C
Tensile Yield Strength	2.8e+008 Pa
Compressive Yield Strength	2.8e+008 Pa
Tensile Ultimate Strength	3.1e+008 Pa

Sources: Authors Compilation

Table-3: Geometrical Properties

Volume	5.6715e-003 m³
Mass	15.71 kg
Nodes	1647
Elements	700
Length	1.5 m

Sources: Authors Compilation

Table-4: Loads

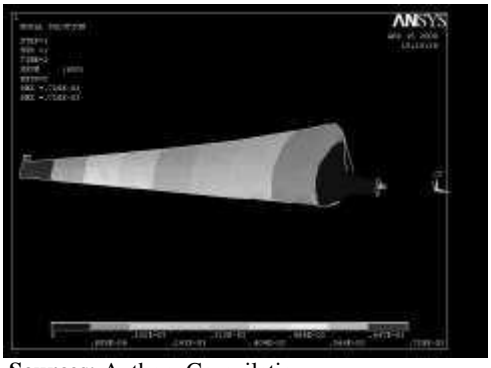
X Component	0.25183 N
Y Component	475.67 N
Z Component	8.4523 N

Sources: Authors Compilation

The static analysis is performed with the above material properties and boundary conditions. The deformations along all the directions and the stresses in those directions are also analyzed. The maximum deformation of 0.728 mm occurs at the tip of the blade. In addition, maximum stress value of 14.9 Mpa occurs nearer the root of the blade. The contour plots for the deformation and stress developed are shown in the following figures:

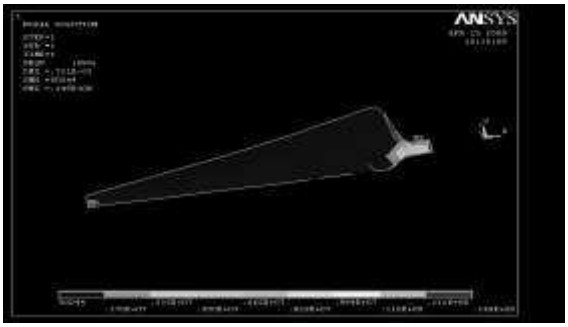
Presentation of Results

Figure-12: Contour Plot for total Deformation



Sources: Authors Compilation

Figure-13: Contour Plot for Von-Mises Stress



CONCLUSIONS

This chapter is the concluding point for this project. This chapter explains the executive summary of the works carried out in the project. India has the ability to generate the power of 45,000 MW. However, with the available technology the power generation is only 3000 MW. Therefore the development in the technology increases the efficiency. This would lead to highly efficient and powerful HAWT in the on coming years

Conclusion from Analysis

Above plots we conclude that, theoretical and computational results are similar. The effect of various parameters like tip speed ratio, wind velocity and rotor radius over the power output and efficiency are studied. The design of the blade is carried out in CATIA V17. The finite element analysis is carried out for structural efficiency using ANSYS 11. The various forces acting on the blade are read from FLUENT into ANSYS. Based on which structural analysis was performed.

REFERENCES

[1] Hansen, A. C., and Butterfield, C. P., (1993). "Aerodynamics of Horizontal- Axis Wind Turbines", Annual Review of Fluid Mechanics, Volume: 25.

[2] Abbott, I. H., and Von Doenhoff, A. E., (1959). "Theory Of Wing Sections, Including A Summary of Airfoil Data", Dover Publications.

[3] Drela, M., (2001). "Xfoil Subsonic Airfoil Development System",Http://Raphael.Mit.Edu/Xfoil.

[4] Timmer, W. A., and Van Rooy, R. P. J. O. M., "Thick Airfoils ForHawts", Journal of Wind Engineering and Industrial Aerodynamics, Volume: 39, 19925.

[5] Rijs, R. P. P.; Jacobs, P., and Smulders, P. T., (1992). "Parameter Study of the Performance of Slow Running Rotors", Journal of Wind Engineering and Industrial Aerodynamics, Volume: 39.

[6] Riegels, F. W., (1961). "Aerofoil Sections; Results from Wind-Tunnel Investigations, Theoretical Foundations", Butterworths.

[7] Munduate, X., and Coton, F. N., (2000). "An Aerodynamic Method For The Preliminary Design of Horizontal Axis Wind Turbines", Journal Of Advanced Engineering Design, Volume: 40, Number: 1.

*****