CYBER CRIME VULNERABILITIES: AWARENESS, IT ACT AND TECHNOLOGY USED FOR PREVENTION

Sudhakar Singh[1], P.K. Khare[2], Prashant Mor[3]

[1] *Research Scholar, Department of Physics and Electronics, RDVV, Jabalpur, M.P., (India)*

[2] *Professor, Department of Physics and Electronics, RDVV, Jabalpur, M.P., (India)*

[3] *Scientific Officer, Department of Physics and Electronics, RDVV, Jabalpur, M.P., (India)*

ABSTRACT

In a digital age, where online communication has become the norm, internet users and governments face increased risks of becoming the targets of cyber attacks. As cyber criminals continue to develop and advance their techniques, they are also shifting their targets focussing less on theft of financial information and more on business espionage and accessing government information. In light of the growth of IT sector in the country, ambitious plans for rapid social transformation and inclusive growth and India's prominent role in the IT global market, providing right kind of focus for secure computing environment and adequate trust and confidence in electronic transactions becomes one of the compelling priorities for the country. This kind of focus enables creation of suitable cyber security eco system in the country, in tune with globally networked environment and at the same time assures its citizens as well the global community about the seriousness of its intentions and ability to act suitably. Terrorists and criminals use information technology to plan and execute their criminal activities. The increase in international interaction and the wide spread usage of IT has facilitated the growth of crime and terrorism. Because of the advanced communication technology people need not be in one country to organize such crime. Hence terrorists and criminals can find security loopholes in the system and can function from unusual locales instead of their country of residence. Most of such crimes have been originating in developing countries. The widespread corruption in these countries fuel these security hacks. The internet has helped fund such crimes by means of fraudulent bank transactions, money transfer etc. In this paper we present various vulnerabilities, which used by attacker to attack system for wrong purpose. We also presented awareness, IT act and technology which used for preventin.

Keywords: Cybercrim, Denial of Service, Financial Crime, Social Network Crime, Password Access, IT Act

I. INTRODUCTION

Cybercrime is a range of illegal digital activities targeted at organisations in order to cause harm. With new mediums of communication, business and societal activities, growth of newer and varied kinds of crime is inevitable. Computers with the aid of the Internet have today become the most dominant medium of communication, information, commerce and entertainment. The Internet is at once several shopping malls, libraries, universities, news paper, television, movie theatre, post office courier service and an extension of

government and business. It is like life in the real world being extended and carried on in another medium that cuts across boundaries, space, time, nationality, citizenship, jurisdiction, sex, sexual orientation, and age. The Internet, with all the benefits of anonymity, reliability and convenience has become an appropriate breeding place for persons interested in making use of the Net for illegal purposes, either monetary or otherwise.

1.1 Computer Wrongs

Computer wrongs includes both civil wrongs and crimes. 'Cyber crime' is used in a generic sense which tends to cover all kinds of civil and criminal wrongs related to a computer. However, the phrase 'cyber crimes' has two limitations to it: (a) 'cyber' generally tends to convey the feeling of 'internet' or being 'online' and hence, does not cover other computer related activities; (b) 'crimes' restricts the application of the phrase to criminal wrongs. It would not include civil wrongs. Thus, it would be preferable to understand the concept of any wrong related to computer as being a 'computer wrong'. It would include any tort or civil wrong done which relates to a computer as also any criminal activity relatable to a computer. One must also keep in mind that it is the statute on a particular subject which informs us as to: (a) whether a particular act is a wrong; and, (b) if it is, whether such wrong is a civil wrong or a crime **[1]**.

1.2 Classification of Computer Crimes

Technology-aided crimes can essentially be classified under two headings:

(i) Where computer is used a tool to commit the crime: The computer is a tool for an unlawful act where the offence reflects a modification of a conventional crime by making use of information technology and modern communication tools.

(ii) Where the computer is the target for the crime: There are certain crimes where the computer itself is the target, that is, to say such crimes which have evolved due to the advancement in information technology itself.

There might be instances where the computer is a tool as well as the target of a crime. This kind of activity involves sophisticated crimes usually out of the purview of conventional criminal law. There is a third category as well, where computers are considered as incidental to a crime. The use of a computer is not necessary but is used to make the offender more efficient in the commission of the crime. This includes use of computers in bookmaking or drug-dealing. The following are some of the most prevalent ways by which hacker can get into a computer system one is physical intrusion, second is remote intrusion and third is remote intrusion **[1,2]**.

Physical intrusion is the most basic of the techniques and most often the most overlooked in information security procedures adopted by IT professionals. If the hacker has physical access such as access to the console or the keyboard then it is very simple for him or her to get into the machine and take the machine apart. The disk may be removed and read/write on another machine. Data can be transferred from the machine to a disk or another machine. With the advent of blue tooth and wireless communication, intrusion has become easier. System intrusion is common where the hacker has access to the system as a low privilege user on the computer system and uses his low privilege account to gain additional administrative privileges. In this scenario the hacker uses security loopholes if the computer system does not have the latest security patches. In remote intrusion the hacker has no physical or user access to the computer system and attempts to hack the computer system remotely across the network**[1,3]**. The network may be an internal company intranet or through the Internet. The main objective of our research work is find out current scenario of cyber crime and different prevention methods from these crime and also present various type of punishment for cyber criminal.

II. RELATED WORK

The increasing online population has proved a happy hunting ground for cyber criminals, with losses due to cyber crime being in billions of dollars worldwide. While other countries are reporting enormous losses to cyber crime, as well as threats to enterprises and critical information infrastructure, there are hardly any such reports coming out of India other than those relating to cyber espionage. Though the report of the National Crime Records Bureau (NCRB) for 2010 reported an increase of 50% in cyber crime over the previous year, the numbers were quite small in absolute terms. The total number of cases registered across various categories was 698, but these low numbers could be because cyber laws have proved ineffective in the face of the complex issues thrown up by Internet. As a case in point, though the cyber crimes unit of the Bengaluru Police receives over 200 complaints every year, statistics show that only 10% have been solved, a majority of these are yet to be even tried in the courts and the cases that did reach the courts are yet to reach a verdict since the perpetrators usually reside in third countries**[4]**. Even though the Information Technology Act (IT Act) 2000 confers extraterritorial jurisdiction on Indian courts and empowers them to take cognisance of offences committed outside India even by foreign nationals provided "that such offence involves a computer, computer system or computer network located in India", this has so far existed only on paper.

Similarly, there are relatively few reports of Indian companies suffering cyber security breaches of the sort reported elsewhere. Companies attribute this to the primacy placed on information assurance in the outsourcing business. Industry bodies such as the National Association of Software and Services Companies (NASSCOM) also attribute this to the fact that they have been at the forefront of spreading information security awareness amongst their constituents, with initiatives such as the establishment of the Data Security Council of India (DSCI) and the National Skills Registry. The Indian government has also aided these initiatives in a variety of ways, including deputing a senior police officer to NASSCOM to work on cyber security issues, keeping the needs of the outsourcing industry in mind. That said, cyberspace is increasingly being used for various criminal activities and different types of cyber crimes, causing huge financial losses to both businesses and individuals. Organized crime mafia have been drawn to cyberspace and this is being reflected in cyber crimes gradually shifting from random attacks to direct (targeted) attacks. A cyber underground economy is flourishing, based on an ecosystem facilitated by exploitation of zero-day vulnerabilities, attack tool kits and botnets. The vast amounts of money lubricating this ecosystem is leading to increased sophistication of malicious codes such as worms and trojans. The creation of sophisticated information stealing malware is facilitated by toolkits such as ZueS, which are sold on Internet for a few thousands of dollars. At the other extreme components of critical infrastructure such as Programmable Logic Control (PLC) and Supervisory Control and Data Acquisition (SCADA) systems were targeted by the Stuxnet malware that attacked supposedly secure Iranian nuclear facilities. Stuxnet exploited five distinct zero-day vulnerabilities in desktop systems apart from vulnerabilities in PLC systems and exposed the grave threat to critical infrastructure such as nuclear plants and other critical infrastructure.

Cyber criminals are using innovative social engineering techniques through spam, phishing and social networking sites to steal sensitive user information to conduct various crimes, ranging from abuse to financial frauds to cyber espionage. While large enterprises are ploughing more resources into digital security, it is the small enterprises and individuals that are falling prey to cyber crime as evinced by the increasing number of complaints on consumer complaint forums. The low levels of computer security are also apparent in recurring statistics that show that India is the third largest generator of spam worldwide accounting for 35% of spam

zombies and 11% of phishing hosts in the Asia-Pacific-Japan region. Over 6,000,000 computers were part of bot Network. India ranked first in the Asia-Pacific region and contributed 21% to the regional total **[1,4]**. A continuing trend for Internet users in India was that of the threat landscape being heavily infested with worms and viruses. The percentage of worms and viruses in India was significantly higher than the Asia-Pacific regional average. According to CERT-In, India sees an average of 788 bot infected computers per day. With regard to web based attacks, India has seen a significant increase and has ranked seventh with 3% of the world attacks and second in the Asia-Pacific region.

III. VULNERABILITIES AND EXPLOITATION OF VULNERABILITIES

Hackers do not magically get into the computer system or information systems, they exploit the technological vulnerabilities present in a computer system, information system or network and then gain access to the computer system. Hacking is a common way of piercing the defences of protected computer systems and interfering with their functioning. Identity theft is also common. The scope and nature of threats and vulnerabilities is multiplying with every passing day. The following paragraphs attempt to provide a brief understanding of the various technological vulnerabilities **[1,5,]**.

3.1 Software Bugs

It is one of the most important ways, which the hackers exploit to gain access into the computer systems. Software bugs can be broadly classified into buffer overflows, unexpected combinations and race conditions. A typical example is a programmer who sets aside 256 characters to hold a login username. However, if an attacker tries to enter in a false username longer than the actual you might have a problem. All the attacker has to do is send 300 characters, including code that will be executed by the server, and thus gain access. Hackers find these bugs in several ways. First, the source code for a lot of services is available on the net. Hackers routinely look through this code searching for programs that have buffer overflow problems. Secondly, hackers may look at the programs themselves to see if such a problem exists. Thirdly, hackers will examine every place the program has input and try to overflow it with random data. If the program crashes, there is a good chance that carefully constructed input will allow the attacker to gain access. Unexpected combinations are scenarios where hackers send input that is meaningless to one layer, but meaningful to another layer. The program is usually constructed using many layers of code and therefore by trial and error method the hacker talks to one of the layers of the software and setting off a chain reaction in other layers, which provides him with the access. Race conditions are scenarios where one program accesses data and the same data is accessed by another program being run by another person which enables the person to access the data. Race conditions work because most computers are designed to handle more than one program at a time. In yet another kind of intrusion, the hacker just feeds random inputs into the system hoping to elicit a response from the system and at times this works.

3.2 System Configuration Bugs

It is security holes, which develop in the system due to the way the system has been configured for use usually by the administrator. Default configurations (configurations in which the system is shipped to the customer) in a system is the most vulnerable and can be hacked in easily. If the administrator fails to set up a root/administrator password in a system it becomes easy for the hacker to gain access. Also in systems, which have been

interconnected with a pool of other systems, then the security loopholes in one unsecure system can be used to hop to other system in the pool, thereby endangering the entire network.

3.3 Internet Browsers and Operating Systems

It also have security holes, which are regularly exploited by hackers to install bugs, viruses and Trojans or for them to be downloaded through various infected sources. This includes URL, HTTP, HTML and JavaScript, Frames, Java and ActiveX attacks. Regular patches are available which need to be used in order to plug these loopholes. The section at the end of this paper provides a list of the most active vulnerabilities, which may be used as a reference. By sending illegal or strange ICMP or TCP packets, a hacker can identify the OS on the target system. Standards usually state how machines should respond to legal packets but omit to instruct the machine how to respond to in valid inputs. Therefore each reply to an invalid input can be used by the hacker to determine and identify the system OS and plan the attack.

3.4 Password Access

Password access is the key to any computer system or in fact networks. Therefore control over password access is perhaps most crucial in ensuring information security and also easiest for the hacker to exploit as a vulnerability. The first major flaw in password access is weak or easy to guess passwords. These passwords are where people use names of pets, loved ones, nick names as passwords thereby enabling the hacker to guess the password easily. Too many passwords are easily guessed, especially if the hacker knows something about their target's background. It's not unusual, for example, for office workers to use the word "password" to enter their office networks. Other commonly used passwords are the computer user's first, last or child's name, secret, names of sports teams or sports terms, and repeated characters such as AAAAAA or bbbbbb. Another method of intrusion exploiting the computer system is dictionary attack on the system. The hacker will use a program, which will try every possible word in the dictionary. Similar to the dictionary attack is the 'brute force' attack where the hacker tries combinations of the password characters in order to break in. A simple five-letter password using English characters may be easy to break in. Sniffing programs on servers or switched networks may prove to be effective in tapping into the user password when he/she logs onto the system. There are other sophisticated methods of gaining password control such as encrypted sniffing and replay attack.

Another interesting mechanism used to gain access to passwords is through social Engineering. 'Social engineering' is hacker speak for conning legitimate computer users into providing useful information that helps the hacker gain unauthorized access to their computer system. Some of the more common social engineering scenarios are.

(i) The attacker pretends to be a legitimate end-user who is new to the system or is simply not very good with computers. The attacker may call systems administrators or other end-users for help. This "user" may have lost his password, or simply can't get logged into the system and needs to access the system urgently. The attacker may sound really lost so as to make the systems administrator feel that he is, for example, helping a damsel in distress. This often makes people go way out of their way to help.

(ii) The attacker pretends to be a VIP in the company, screaming at administrators to get what he wants. In such cases, the administrator (or it could be an end-user) may feel threatened by the caller's authority and give in to the demands.

(iii) The attacker advantage of a system problem that has come to his attention, such as a recently publicized security vulnerability in new software. The attacker gains the user's trust by posing as a system

administrator of maintenance technician offering help. Most computer users are under the mistaken impression that it is okay to reveal their password to computer technicians.

(iv) The attacker posing as a system administrator or maintenance technician can sometimes persuade a computer user to type in computer commands that the user does not understand. Such commands may damage the system or create a hole in the security system that allows the attacker to enter the system at a later time.

3.5 Insecure Modems

It is another gateway for a hacker to gain access to a computer system. War dialers are used by hackers to identify the modems of a target. A war-dialer is a computer program that automatically dials phone numbers within a specified range of numbers and chances are that if an organization has one number, it will have a few other numbers in same range for all telecommunications. By dialing fall numbers within the targeted range, the war-dialer identifies which numbers are for computer modems and determines certain characteristics of those modems. The hacker then uses other tools to attack the modem to gain access to the computer network. Effective war-dialers can be downloaded from the internet at no cost. The problem is that modem is a means of bypassing the "firewall" that protects your network from outside intruders. A hacker using a "war-Dialer" to identify the modem telephone number and a password cracker to break one weak password can gain access to the system. Due to the nature of computer networking, once a hacker connects to that one computer, the hacker can often connect to just about any other computer in the network. Of course it is now possible to incorporate safeguards to prevent easy access through modems, which is beyond the scope of this paper.

3.6 Cookies

It is another security threat that the user of a computer system faces. A cookie is a small program that may be placed on a computer. The cookie enables the site that has deposited the cookie to recognise when the user visits it the next time. It maintains a database of the users visits to the site and also in some instances other websites. Cookies raise substantial privacy issues, which are again beyond the scope of this paper. Suffice to say that cookies do raise issues of profiling of individuals, illegal tracking on the Internet etc. Cookies per se do not damage or hack the system but are often used by hackers to gain information on a target and his/her Internet surfing habits prior to hacking. It is possible to ensure that the user's computer systems do not accept cookies from any site and settings on the system and special software installation will achieve this goal.

3.7 Denial of Service

This attacks are another variety of system compromises which are designed to overload network links, the processing unit of the user system or the disk of the system thereby crashing the service. The hacker aims to make the computer system deny providing services to the user. The increased degree of automation in the recent years has enabled a single hacker to control thousands of compromised systems for use in the attacks. A simple example may be to flood the user's (in most case an entire organization's) mail inbox with a host of messages thereby making the server to crash. In the recent past attacks on internet domain name system (DNS) is on the rise. The hacker may create a bogus DNS resembling a legitimate internet site. Therefore information intended for the legitimate site flow into the hacker's site. In some other cases hackers compromise poorly protected DNS servers which give them the ability to modify the data passing through the server. By leveraging insecure mechanisms used by customers to update their domain registration information, attackers can co-opt the domain registration processes to take control of legitimate domains. Another issue which has cropped up recently is web

spoofing which is a kind of electronic con game in which the attacker creates a convincing but false copy of the entire world wide web. The false web looks just like the real one: it has all the same pages and lings. However, the attacker controls the false web, so that all the network traffic between the victim's browser and the web goes through the attacker. The key to this attack is for the attacker's web server to sit between the victim and the rest of the web. This kind of arrangement is called a 'man in the middle attack' in the security literature. Since the attacker can observe or modify any data going from the victim to web servers, as well as controlling all return traffic from web servers to the victim, the attacker has many possibilities. These include surveillance and tampering.

3.8 Attacks against Routers

It is another vulnerability that may be exploited by hackers to crash information systems. Intruders use poorly secured routers as platforms for generating attack traffic at other sites, or for scanning or reconnaissance. Further, routers are designed to pass large amounts of traffic through them; they often are not capable of handling the same amount of traffic directed at them. Intruders take advantage of this characteristic attacking the routers that lead into a network rather than attacking the systems on the network directly. Another method of intrusion into routers is to exploit the trust relationships that the routers have. For routers to do their job, they have to know where to send the traffic they receive. They do this by sharing routing information between them, which requires the routers to trust the information they receive from their peers. As a result, it would be relatively easy for an attacker to modify, delete, or inject routes into the global internet routing tables to redirect traffic destined fro one network to another, effectively causing a denial of service to both (one because no traffic is being routed to them, and the other because they're getting more traffic than they should). Although the technology has been widely available for some time, many networks (Internet service providers and large corporations) do not protect themselves with the strong encryption and authentication features available on the routers. Table 1: shows various sectors prone to cyber attacks**[6]**. Figure 1: shows chart of cyber attack in different sectors. Financial services sector is most prone to cybercrime.

Table 1: Sectors Prone To Cyber Attacks

Sn.	Sectors	Rate of Cyber attacks
1	Consumer/Industrial Markets	3%
2	Pharmaceuticals	3%
3	Infrastructure	11%
4	Energy and Natural resources	3%
5	Communication/Entertainment	11%
6	Financial services	58%
7	Government	8%
8	Others	5%

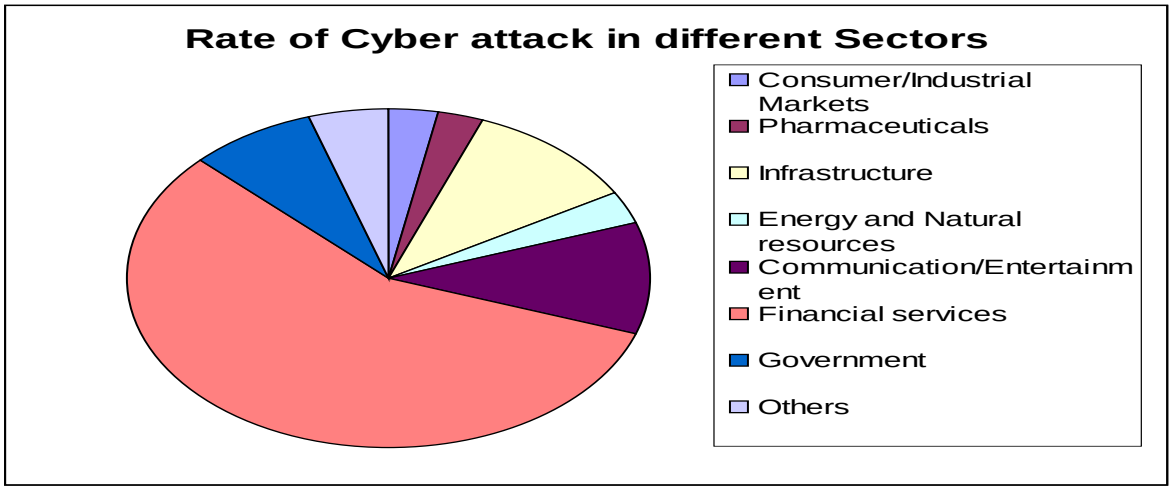


Figure 1: Rate of Cyber Attack in Different Sector

IV. INVESTIGATION OF CYBER CRIMES

4.1 GPS Application

The GPS system concept is based on *time*. The satellites carry very stable atomic clocks that are synchronized to each other and to ground clocks. Any drift from true time maintained on the ground is corrected daily. Likewise, the satellite locations are monitored precisely. GPS receivers have clocks as well—however, they are not synchronized with true time, and are less stable. GPS satellites continuously transmit their current time and position. A GPS receiver monitors multiple satellites and solves equations to determine the exact position of the receiver and its deviation from true time. At a minimum, four satellites must be in view of the receiver for it to compute four unknown quantities (three position coordinates and clock deviation from satellite time). GPS has become a widely deployed and useful tool for commerce, scientific uses, tracking, and surveillance. GPS's accurate time facilitates everyday activities such as banking, mobile phone operations, and even the control of power grids by allowing well synchronized hand-off switching **[7,8]**.

4.2 Digital Intelligence Collection

The value of collecting intelligence informatin about threat sources and possible cyber attacks cannot be underestimated. A well deployed cyber attack can yield vital information that compromises communication and encryption ciphers. It tends to project the power of the attacker and demoralize the victim. However, the changing phase of cyber attacks as well as ever-increasing sophistication of attack methods have complicated the efforts of collecting valuable intelligence information for effective proactive, preventive and protective measures. Generally, attacks directed against Govt. and critical information infrastructure can be categorized as either massive attacks, aimed at disabling the infrastructure rendering it unusable or inaccessible to users; or targeted attacks, aimed at collecting sensitive/strategic information. Massive attacks generally take the form of denial of service attacks against the infrastructure. The targeted attacks involve a good deal of customization and personalization of attack methods and levels of technological and operational sophistication. Skillful execution of attack and the methodology used to conceal any traces of attack complicates the task of advance intelligence information collection and/or attack detection**[9].**

4.3 CDR Analysis

CDR details for the specified period including cell tower details.

4.4 Email Testing

The Forensic investigation of the network and e-mails revealed a list of IP (Internet Protocol) addresses which were linked to the perpetrator. Analysis showed that the perpetrator succeeded in creating and uploading several files to the e-mail server through a malicious code, allowing the perpetrator to send commands to the server from a remote location and gather sensitive business information. Reconstruction of the malicious code installation revealed it was executed by an internal employee who in hand with the perpetrator managed to extract sensitive information.

4.5 Social Network Crime Management

Social engineering is the act of manipulating people into performing actions or divulging confidential information, rather than by breaking in or using technical cracking techniques. This method of deception is commonly used by individuals attempting to break into computer systems, by posing as an authoritative or trusted party and capturing access information from the naive target. Email Phishing is a common example of social engineering's application, but it is not limited to this single type of attack some are given in the following table **[2]**.

Table2: Various Threats Related To Communities

Sn.	Threats to Communities
1	Propagation of malware through social networks / Web navigation
2	Cyber-extortion / demand for ransom / blackmail
3	Loss or theft of personal and confidential data
4	Misinformation / political actions / damaged image
5	Denial of service / blocked access / paralysis / unavailability

Multimedia data, such as voice, text and network session data, is compiled and processed. Through this compilation and processing, names, entities, relationships and individual events are extracted from the multimedia data. This information is then used to perform a social network analysis on the criminal network, through which the user can detect and track threats in the network.

4.6 Financial Crime Management

Financial crime can expose a financial institution to various risks, including operational, legal, regulatory and reputational risks. The extent of these risks, alone or in combination, is particularly wide when the perpetrators take advantage of deficiencies in the institutions management or the complicity of its employees or officers. The financial institution should manage financial crime risks within its integrated risk management framework. Accordingly it should give consideration to the interrelationships and interdependencies between risks. This means that the institution have following.

(i) identify, assess and quantify financial crime risks;

(ii) implement risk mitigation measures in order to reduce the likelihood of events that could affect the institution.

4.7 Information Technology ACT, 2000

The Information Technology Act intends to give legal recognition to e-commerce and e-governance and facilitate its development as an alternate to paper based traditional methods. The Act seeks to protect a common man from the ill effects of the advancement in technology by defining crimes, prescribing punishments, laying down procedures for investigation and appointing regulatory authorities. Many electronic crimes have been bought within the definition of traditional crimes too by means of amendment to the Indian penal code, 1860. The Evidence Act 1872 and the Banker's Book Evidence Act, 1891 too have been suitably amended in order to facilitate collection of evidence in fighting electronic crimes.

In the following common computer crimes are described and also included under the Indian Information Technology Act, 2000 the Indian penal code and other minor criminal Acts have been described. The computer crimes can be classified into the following categories.

(i) Conventional crimes through computer: cyber defamation, digital forgery, cyber pornography, cyber stalking/harassment, Internet fraud, financial crimes, online gambling, and sale of illegal articles.

(ii) Crimes, committed on a computer network: hacking/unauthorized access, denial of service.

(iii) Crimes relating to data alteration/destruction: virus/worms/Trojan horses/logic bomb, theft of Internet hours, data diddling, salami attacks,

4.7.1 Offences under the IT Act

Chapter XI of the act enumerates the various acts which constitute an offence under the act along with the punishment be it either imprisonment or fine or both. Such offences (including under other sections) can be better understood in the form of a table 3 **[10, 11].**

Table3: Various Section of IT Act 2000

Section	Offence	Punishment
33(2)	Failure of any Certifying Authority to surrender a licence under Section 33(1) after such licence has been suspended or revoked [Section 25(1)].	Person in whose favour the licence is issued shall be punished with imprisonment which may extend up to six months or a fine which may extend up to Rs. 10,000 or both.
65 (Tampering)	Knowingly or intentionally concealing, destroying or altering or intentionally or knowingly causing another to conceal destroy, destroy, or alter any computer source code use for a computer, computer program ,computer system or computer network, when the computer source code is required to be kept or maintained by law for the time being in force.	Punishable with imprisonment up to three years, or with fine which may extend up to Rs. 2,00,000/-, or with both.
66(Hacking)	Destroying or deleting or altering any information residing in a computer resource or diminishing its value or utility	Punishable with imprisonment up to three years, or with fine which may extend up to Rs. 2,00,000/-, or with both.

	or affecting it injuriously be any means with the intent to cause or knowing that he is likely to cause wrongful loss or damages to the public or any person.	
67	Publishing or transmitting or causing to be published in the electronic form, any material which is lascivious or appeals to the prurient interest or if its effect is such as to tend to deprave and corrupt persons who are likely, having regard to all relevant circumstances, or read, see or hear the matter contained or embodies in it that is hacking as defined under Section 67(1)	First conviction: punishable with imprisonment of either description of a term which may extend to five years and with fine which may extend to Rs. 1,00,000/-, Second or subsequent conviction: imprisonment of either description of a term which may extend to ten years and with fine which may extend to Rs. 2,00,000/-.
68(2)	Failure to comply with the order of controller under section 68(1) which empowers the controller to direct, by order, a certifying Authority or any employee of such authority to take such measures or cease carrying on such activities as specified in the order if those are necessary to ensure compliance with the provisions of this Act, rule or any regulation made there under.	Punishable with imprisonment for a term not exceeding three years or to a fine not exceeding Rs. 2,00,000/-, or to both.
69(3)	Failure to assist an agency [referred in section 69(2)] which is required to intercept any information as required by an order of the controller [under section 69(1)	Punishable with imprisonment for a term which may extend to seven years.
70(3)	Securing access or attempting to secure access to a protected system [as declared by the appropriate Government vide a notification under section 70 (1)] in contravention of the provisions of this section [that is such person is not authorized by the appropriate Government under section 70(2) to access the protected system].	Punishable with imprisonment of either description for a term which may extend to ten years and shall also be liable to fine.
71	Making any misrepresentation to, or suppressing any material fact from, the	Punishable with imprisonment for a term which may extend to two years, or with

	Controller or the certifying Authority for obtaining any licence or digital signature certificate.	fine which may extend to Rs. 1,00,000/- or with both.
72	Securing access to any electronic record, book, register, correspondence, information, document or other material by any person in pursuance of any of the powers conferred under this Act, rules or regulations made thereunder without the consent of the person concerned and thereafter, disclosing such electronic record, etc. to any other person.	Punishable with imprisonment for a term which may extend to two years, or with fine which may extend to one Rs. 1,00,000/- or with both.
73	Publishing a digital signature certificate or otherwise making it available to any other person with the knowledge that- (a) the certifying authority listed in the certificate has not issued it; or, (b) the subscriber listed in the certificate has not accepted it, or, (c) the certificate has been revoked or suspended, unless such publication is for the purpose of verifying a digital signature created prior to such suspension or revocation.	Punishable with imprisonment for a term which may extend to two years, or with fine which may extend to one lakh rupees or with both.
74	Knowingly creating, publishing or otherwise making available a digital signature certificate for any fraudulent or unlawful purpose.	Punishable with imprisonment for a term which may extend to two years, or with fine which may extend to one lakh rupees, or with both.

In case of offences committed by companies, such persons who, at the time the contravention was committed, was in charge of, and was responsible, to the company for the conduct of business of the company as well as the company, will be, under sub-section (1) of section 85 of the Act, guilty of the contravention and shall be liable to be proceeded against and punished accordingly. However, if such personal proves that the contravention took place without his knowledge or that he exercised all due diligence to prevent such contravention, he shall not be liable to punishment. Sub-section (2) of section 85 also deems a director, manager, secretary or any other office of the company to be guilty of contravention and liable for punishment if it is proved that the contravention has taken place with the consent or connivance of, or is attributable to any neglect on the part of such person. 'Company', for the purpose of this section, has been explained to mean any body corporate and includes a firm, or other association of individuals. 'Director', in relation to a firm, would mean a partner in the firm. Section 7 prohibits immunity against any punishment under any other law to which a person might be liable to in spite of any penalty imposed or confiscation made under the IT Act.

4.7.2 Investigation Under the IT Act

The procedure for investigation for computer crimes is no different from the investigation for conventional crimes and code of criminal procedure, subject to the provisions of the IT Act, would apply. Investigation, for the purposes of the code of criminal procedure, 1973, has been held by the supreme court [State of Maharashtra v. Rajendra, (1997) 3 crimes 285] to consist generally of the following steps:

1) Proceeding to the spot
2) Ascertaining all the facts and circumstances of the case
3) Discovery and arrest of the suspected offender
4) Collection of evidence relating to the commission of the offence which may consist of, the examination of various persons (including, the accused) and the reduction of their statement into writing if the officer things fit, the search of places and seizure of things considered necessary for the investigation and to be produced at the trial, and Formation of the opinion as to whether on the materials collected, there is a case to place the accused before a magistrate for trial and if so, taking the necessary steps for the same by filing a charge-sheet under section 173.

Section 78 of the IT Act places the powers of investigation to a police officer not below the rank of Deputy Superintendent of police. This provision overrides anything contrary in the code of criminal procedure. Section 80 enumerates the powers of police officers to enter and search premises. Sub-section (1) of section 80 provides that any police officer, not below the rank of a Deputy superintendent of police, or any other officer of the central government or a state government authorized by the central government in this behalf may enter any public place and arrest without warrant any person found therein who is reasonably suspected of having committed or of committing or of being about to commit any offence under the Act. For the purposes of sub-section (1) the expression 'public place' has been explained to include any conveyance any hotel, any shop or any other place intended for use by or accessible by the public. Where any person is arrested under sub-section (1), then sub-section (2) requires that such person should, without unnecessary delay, is taken or sent before a magistrate having jurisdiction in the case or before the officer-in-charge of a police station. The provisions of the code of criminal procedure are to apply in relation to any entry, search or arrest made under section 80, subject of course to the provisions of the section itself. Time to time IT act amendment is added for handle cyber crime related problem. For example Information Technology (Amendment) Act 2008.

V. RESULTS AND DISCUSSION

A cyber countermeasure is defined as an action, process, technology, device or system that serves to prevent or mitigate the effects of a cyber attack against a computer, server, network or associated device. Recently there has been an increase in the number of international cyber attacks. In 2013 there was a 91% increase in targeted attack campaigns and a 62% increase in security breaches **[12,13].**

5.1 Security Awareness

Many cyber vulnerabilities exist because of lack of information security awareness on the part of computer users, system or network administrators, technology developers, auditors, Chief Information Officers (CIOs), Chief Executive Officers (CEOs) and Corporates. A lack of trained personnel and the absence of widely accepted, multi-level certification programs for information security professionals complicate the task of

addressing cyber vulnerabilities. This policy identifies following major actions and initiatives for user awareness, education and training**[4]**.

(i) Promoting a comprehensive national awareness program

(ii) Fostering adequate training and education programs to support the Nation's information security needs (Ex School, college and post graduate programs on IT security)

(iii) Increase in the efficiency of existing information security training programs and devise domain specific training programs (ex: Law Enforcement, Judiciary, E-Governance etc)

(iv) Promoting private sector support for well coordinated, widely recognized professional information security certifications.

The main purpose is to achieve the broadest penetration to enhance awareness and alert larger cyber community in cases of significant security threats. The promotion and publicity campaign could include

(i) Seminars, exhibitions, contests etc

(ii) Radio and TV programmes

(iii) Videos on specific topics

(iv) Web casts and Pod casts

(v) Leaflets and Posters

(vi) Suggestion and Award Schemes

5.2 Technical Control by Services Provider

ISP (Internet Services Provider) with play a significant role in building trust in online transactions and making the best use of multifaceted technology. The character of cyber crime is such that even though the offender acts alone, several entities get automatically involved. For instance, sending a simple email requires the service of the e-mail provider, access providers and the routers who forward the e-mail message to the recipient. In achieving this objective these organizations are faced with the dilemma of how best to collaborate with each other to make the Internet safer without infringing the fundamental rights of users. The role of ISP is essential as it is not possible to commit cyber crime without their involvement in one way or another. But it is not always possible for the ISP to prevent the commission of cyber crime. The point to be looked into is whether their role should be limited. The implications of the answer directly affect the economic development of the ICT infrastructure. The ISP must be closely associated in providing for secure information flow through their Network and gateways..

The role and legal responsibility of the Network service providers in India has been defined by the IT Act. Section 79 restricts the liability of service providers in certain cases. It provides that the service providers shall not be responsible for any third party information or data provided by them if they had no knowledge of it or had exercised all due diligence to prevent the offence from being committed. However, these provisions shall not apply if the service provider has conspired, abetted or aided in the commission of the offence or where the ISP fails to cooperate with the government in preventing the commission of an unlawful act. Their liability in India is also determined by Licence for Internet Services, Clause 33 and Clause 34 of which set out the various responsibilities of the service providers, some of which are given below **[4,9-11]**.

• ISP must prevent unlawful content, messages or communications including objectionable, obscene, unauthorized content from being carried on their Network.

• They must ensure that content carried by them does not infringe cyber laws.

• They must comply with the IT Act provisions and must assist the government in countering espionage, subversive acts,
sabotage or any other unlawful activity.
• Privacy of communication online is ensured by preventing unauthorized interception of messages.
• Government can take over their equipment and Network in times of emergency war etc.

5.3 Actions by Large Corporates

• Compliance to international security best practices and demonstration
• Pro-active actions to deal with and contain malicious activities, and protecting average end users by say of net traffic monitoring, routing and gateway controls
• Keeping pace with changes in security technology and processes to remain current (configuration, patch and vulnerability management)
• Conform to legal obligations and cooperate with law enforcement activities including prompt actions on alert/advisories issued by CERT-In
• Use of secure product and services and skilled manpower
• Crisis management and emergency response.
• Periodic training and up gradation of skills for personnel engaged in security related activities
• Promote acceptable users' behavior in the interest of safe computing both within and outside.

5.4 Actions by Small/Medium Users and Home Users

• Maintain a level of awareness necessary for self-protection
• Use legal software and update at regular intervals.
• Beware of security pitfalls while on the net and adhere to security advisories as necessary
• Maintain reasonable and trust-worthy access control to prevent abuse of computer resources

VI. CONCLUSION

The IT (Information Technology) sector has become one of the most significant growth catalysts for the Indian economy. In addition to fuelling India's economy, this sector is also positively influencing the lives of its people through direct and indirect contribution to the various socio-economic parameters such as employment, standard of living and diversity among others. The sector has played a significant role in transforming India's image to that of a global player in providing world class technology solutions and business services. The government has been a key driver for increased adoption of IT based products and solutions in the country. It has embarked on various IT enabled initiatives including in Public services (Government to citizen services, citizen identification, public distribution systems), Healthcare (telemedicine, remote consultation, mobile clinics), Education (e-Learning, virtual classrooms) and Financial service (mobile banking/payment gateways), etc. In addition, Government sector has enabled increased IT adoption in the country through sectors reforms that encourage IT acceptance and National programmes such as National eGovernance Programmes (NeGP) and the Unique Identification Development Authority of India (UIDAI) programme that create large scale IT infrastructure and promote corporate participation. In conclusion we see that today cyber crime problem faced by almost all sector, because computer or similar devices used in every sectors. Therefore to handle these problem combined approach like security awareness, technical control by Services Provider, actions by Large Corporates, actions by small/medium users and home users , uses of IT act etc. are required.

REFERENCES

[1] Introduction of Cyber crime; IGNOU study material of PGCCL program; MIR 012 BLOCK 02, March 2008

[2] Singh Brijendra., "Network Security and Management", Prentice Hall of India Private Limited, New Delhi-110001, Published in 2007.

[3] Stalling, William., "Network Security Essentials application and standards", Third Edition, Pearson Prentice Hall, Published in 2008

[4] Report of Institute for Defence Studies and Analyses, New Delhi ; Published in March 2012 ; http://www.idsa.in

[5] FFIEC Information Technology Examination Handbook, "Operations" ; www.ithandbook.ffiec.gov/it-booklets/operations.aspx

[6] KPMG Cybercrime survey report in India; Year 2014

[7] "Global Positioning System"; www. Gps.gov; Retrieved Feb, 2015

[8] http://en.wikipedia.org/wiki/Global_Positioning_System

[9] Department of Information Technology Ministry of Communications and Information Technology Government of India; National Cyber Security Policy, draft v1.0, 26 Mar 2011

[10] Information Technology Act. 2000; PGCCL, IGNOU study material of cybercrime; March 2008

[11] Information Technology Act. 2000, The Gazette of India, New Delhi, 9 June 2000

[12] Coleman, Kevin; "Cyber Attacks on Supply Chain Systems" ; defensetech.org; Retrieved 2 May 2011

[13] "Why the US Needs More Cyber Professionals"; Norwich University; Retrieved 23 October 2014.

STUDY OF OPTICAL PROPERTIES OF LITHIUM LEAD SILICATE GLASS

Neetu Ahlawat[1*], Navneet Ahlawat[2]

[1]*Guru Jambheshwar University of Science & Technology, Hisar,(India)*

[2]*Matu Ram Institute of Engineering and Management, Rohtak, (India)*

ABSTRACT

Lithium lead silicate glass system with composition $30Li_2O.xPbO.(70-x)SiO_2$ (where, x= 10, 20, 30, 40, and 50 mol%) were synthesized at 1423K for 0.5 h using normal melt quench technique. The optical transmission spectra within the wavelength range 300-3300 nm of these prepared samples were recorded at room temperature to study their optical properties. It was found that cutoff wavelength decreases as the content of SiO_2 increases in the glass matrix. The variation in optical band gap energy with SiO_2 content may be attributed to influence of SiO_2 on band gap. The values of either direct or indirect optical band gap lies in the range 2.61eV-3.30 eV suggest the semiconducting behavior of these studied glasses. The Urbach energy values calculated for each glass sample shows highest defect concentration for glass sample with x=50mol%. The oxide electronic polarisability of all studied glasses has been calculated by using E_g values. The successive replacement of PbO by SiO_2, results in lower polarisability values in the studied glasses. The low optical basicity means a reduced electron donor ability of oxide ions. Therefore, decrease in optical basicity with PbO/SiO_2 ratio causes a shifting from ionic to more covalent character of the bonding between cation and and oxide ion in these glasses.

Keywords: Semiconductor Glasses, Optical Properties, Heavy metal oxide glasses, Electronic polarisability, Optical basicity,

I. INTRODUCTION

The discovery of heavy metal oxide (PbO, Bi_2O_3, CdO) glasses have attracted interest because of their high non-resonant optical nonlinearity and their good infrared transmittance up to about 7□ B[illegible]3]. Further, these glasses possess a high third order non-linear optical susceptibility, which makes them useful for ultra fast optical switches and photonic devices [4]. These glasses are also promising materials for technological applications such as upconverting phosphors, new laser materials and optical waveguides due to their low phonon energy [5–7]. Silicate glasses with proper amount of lead oxide have the possibility of forming surface layers with high electron surface conductivity and a large electron secondary emission coefficient which makes them promising materials for the production of channeltrons and microchannel plates. Lead silicate system has great importance due to their numerous industrial applications in optoelectronics, radiation shielding etc. [8]. Moreover, the small activation energy of the electric conductivity and negligible changes in other properties within the surface layers mean that they can be applied in a very wide temperature range [9]. Further, photo induced refractive index change in lead silicate glass is thermally stable and strongly correlated with PbO composition. These large thermally stable photosensitivities are attractive for applications in telecommunications, integrated optics and data storage [10]. The properties of glass system depend upon their composition and to a considerable extent

upon their structure. The objective of present work is to study the effect of conventional network former SiO_2, on the optical properties of lithium lead-silicate glasses.

II. EXPERIMENTAL DETAILS

The lithium lead silicate glasses having compositions $30Li_2O\cdot(70\text{-}x)PbO\cdot xSiO_2$ ($10 \leq x \leq 50$ mol %) (LPS glass) were prepared by simple melt-quench technique. Analytical grade chemicals of Li_2CO_3, PbO and SiO_2 were taken in appropriate amounts and thoroughly mixed to form a batch of 20 g. The well-mixed chemicals were then melted at 1423 K in porcelain crucibles using an electrical muffle furnace for 0.5 h. The mixture was stirred frequently to ensure homogeneity. The melt was rapidly quenched in between two stainless steel plates at room temperature (RT) and coin-shaped glass samples were formed. Color of the glasses ranged from dark brown to light yellow when the SiO_2 molar concentration is gradually increased from x=10 to 50 mol%. Glass samples were polished to optical quality for optical measurements. The sample thickness was in the range 0.5–1.5mm. The transmission spectra of the glasses were recorded with a UV/VIS/NIR spectrophotometer (Cary 5000) in the spectral range 200–3300 nm. The cut-off wavelength (λ_{cutoff}) and optical band gap (E_g) were determined from the transmission spectra.

III. RESULTS AND DISCUSSION

The optical transmission spectra of the prepared samples were recorded at room temperature and Fig. 1 shows the optical transmission spectra of all the LPS glasses in the UV-VIS region within the wavelength range 300-800 nm to have a better idea about λ_{cutoff}. The broad absorption edge obtained in Fig. 1 indicates the amorphous nature of LPS

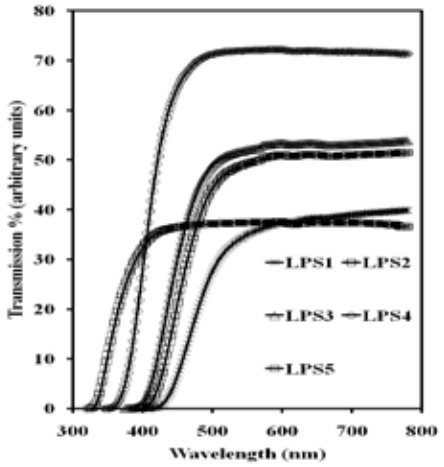


Fig.1 Optical Transmission Spectra for All LPS Glasses

glasses. The values of λ_{cutoff} in nm are given in Table 1. It was found that λ_{cutoff} values decrease as the content of SiO_2 increase in the glass system. The optical band gap in amorphous system is closely related to the energy gap between the valence band and conduction band [11, 12]. The value of optical band gap energy (E_g) can be determined by using the relation [12]:

$$E_g = \frac{1239.4}{\lambda_{cutoff}\,(nm)} \qquad (1)$$

The value of E_g for all LPS glasses as obtained using Eq. (1) are given in Table 1. The variations in E_g values with the SiO_2 content may be attributed to influence of SiO_2 on the band gap [13]. The absorption coefficient

$$a(n) = \frac{B(hn - E_g)^r}{hn} \qquad (2)$$

For amorphous materials, direct and indirect transitions (r = 1/2, 2) are valid according to the Tauc's relations [14]. The corresponding Tauc's plots, {(vs. h) vs. h {(h) vs. and {(h) for LPS glasses are shown in Fig. 2. The values of optical band gap (E_g) for direct and indirect transitions are obtained by extrapolating the linear region of the curves in Fig. 2 to meet h = 0, i(at (= 0) and (h) = 0) respectively [15, 16] and are given in Table 1. The increase in $E_g^{direct/indirect}$ with increasing SiO_2 content for both types of transitions (Table 1) is in consonance with the blue shift in λ_{cutoff}. Fig. 3 shows the variation between Ln(vs h and the

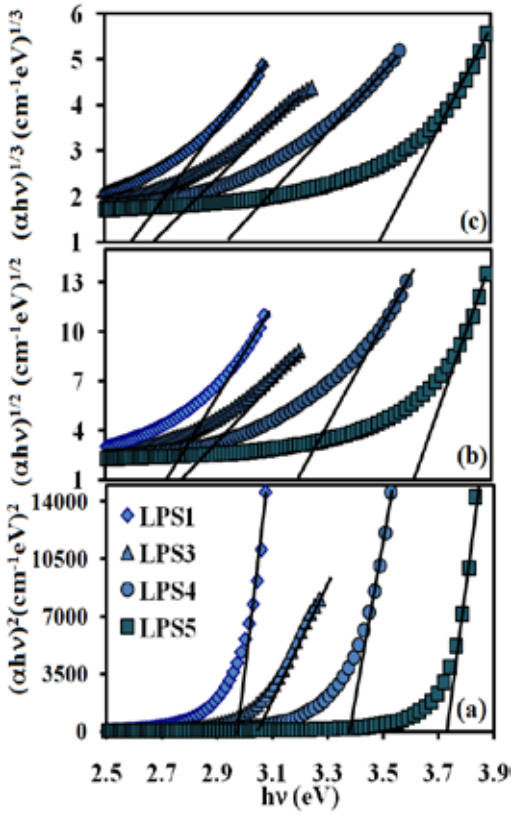


Fig.2 Tauc's Plots For All LSP Glass Samples for (A) R=1/2 (B) R =2 (C) R = 3.

Table1 Composition parameter (x), Cutoff wavelength (λ_{cutoff}), optical band gap (E_g), Urbach energy (E_U), direct band gap (E_g^{direct}), Indirect Band Gap ($E_g^{indirect}$), oxide ion polarisability, theoretical optical basicity (Λ_{th}) and metallisation criterion (M) for $30Li_2O·(70-x)$ PbO·$xSiO_2$ (LPS) glasses.

Glass code	LPS1	LPS2	LPS3	LPS4	LPS5
x(mol%)	10	20	30	40	50
λ_{cutoff} (nm)	414	396	384	352	326
E_g (eV)	2.99	3.13	3.23	3.52	3.80
E_U (eV)	0.26	0.17	0.25	0.27	0.43

E_g^{direct} (eV) for r =1/2	2.99	3.05	3.05	3.30	3.20
($E_g^{indirect}$) (eV)	2.66	2.70	2.61	2.85	3.30
$(\alpha_{O^{2-}})$ ($\times 10^{-24} cm^3$)	3.92	4.26	3.15	3.00	2. 63
Λ_{th}	1.25	1.28	1.14	1.11	1.04
M	0.40	0.36	0.33	0.25	0. 21

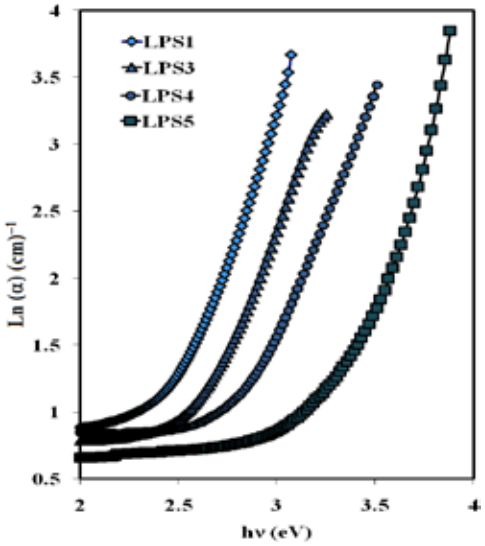


Fig. 3 Urbach's Plots For Some LPS Glass Samples.

values of Urbach energy (E_U) are calculated from the reciprocal of the slopes [17] of the linear portion of these curves for all the LPS glasses [17] and are presented in Table 1. Materials with large Urbach energy would have greater tendency to convert weak bonds into defects. Consequently, the defect concentration could be decided by the measure of Urbach energy. The maximum defect concentration was found in LPS5 glass sample with highest value of Urbach energy. The values of E_U are identical to the experimental observations of defect density and Urbach energy reported in the literature, which shows similar behavior [18]. Dimitrov and Sakka used Duffy's model to derive the relationship between electronic oxide ion polarizability $(\alpha_{O^{2-}})$ and E_g [19, 20] and was later on modified by Banu et al [21]. This relationship has been found to be applicable for many glass systems and is given as:

$$\alpha_{O^{2-}} = \left(\frac{E_g}{\left[\frac{V_M}{2.52}\left\{1 - \frac{(E_g)^{\frac{1}{2}} - 1.14}{0.98}\right\} - \sum_i p_i \alpha_i\right] q}\right) - 1 \qquad (3)$$

where p and q denote the number of cations and oxide ions, respectively in the chemical formula ApOq.The oxide electronic polarisability $\alpha_{O^{2-}}$ of all LPS glasses have been calculated by substituting E_g values (as obtained from Eq. (1)) in this Eq. (3), and are presented in Table 1. It was observed that $\alpha_{O^{2-}}$ decreases with decrease in PbO/ SiO_2 ratio which is an expected result due to the replacement of a highly polarizable lead oxide with Pb^{2+} ion having a lone pair in the valence shell by a less polarizable silicon oxide [22]. According to

Gillespie, [23] repulsion works between the different electron pairs in the coordination polyhedral of the cation. It is known that the bond-pair–lone-pair repulsion is smaller than bond-pair–bond-pair repulsion. This may lead to an additional relaxation of the cation polarizing effect on the oxide ion and the distance Pb-O becomes larger which increases the polarizing effect of Pb^{2+} cation on oxide ions. The successive replacement of PbO by SiO_2, results in lower polarisability values in the studied glasses. The values of theoretical optical basicity (Λ_{th}) can be calculated using oxide ion polarizability $a_{O^{2-}}(E_g)$ obtained from optical band gap by the relation as suggested by Duffy [24]:

$$\Lambda_{th} = 1.67\left(1 - \frac{1}{a_{O^{2-}}}\right) \quad (4)$$

The theoretical optical basicity of all the LPS glasses has also been obtained using the Eq. (4) and is given in Table 1. PbO is an oxide with a significant basicity (0.98) and SiO_2 is a strong acidic oxide with low optical basicity (0.48). It was observed that lead silicate glasses with a large amount of SiO_2 (x=50 mol %) content possess low optical basicity. Low optical basicity means low electron donor ability of the oxide ions. Therefore, decrease in optical basicity with PbO/SiO_2 ratio causes a shifting from ionic to more covalent character of the bonding between cation and oxide ion in these glasses. The necessary and sufficient condition for predicting the non-metallic nature of solids is $R_m/V_m < 1$. The difference from $R_m/V_m = 1$, i.e. M, is so called metallization criterion as given below:

$$M = 1 - \left(\frac{R_M}{V_M}\right) \quad (5)$$

The metallization criterion, M was calculated from above relation is given in Table 1. It was found that 'M' decreases with increasing SiO_2 content, meaning that the tendency for metallization in the electronic structure is high in the glasses with high SiO_2 contents.

IV. CONCLUSIONS

The values of various optical parameters viz. cutoff wavelength, optical band gap energy; optical polarisability, theoretical optical basicity and metallization criterion were calculated from the UV-Vis-NIR spectra recorded for all prepared lithium lead silicate glasses. It was found that these optical parameters can be tailored with the two glass forming anion concentration ratio in the prepared glass compositions. It was concluded that decrease in PbO/SiO_2 ratio causes a shifting from ionic to more covalent character of the bonding between cation and and oxide ion in these glasses.

REFERENCES

[1] W.H. Dumbaugh, Phys. Chem. Glasses **19**, 121 (1978).

[2] J.C. Lapp, W.H. Dumbaugh and M.L. Powley, Riv. Staz. Sper. Vitro. **1,** 91 (1989).

[3] J.E. Shelby, Lead galliate glasses, J. Am. Ceram. Soc.**71,** 254-256 (1988).

[4] A. Pan and A. Ghosh, A new family of lead bismuthate glass with a large transmitting window, J. Non-Cryst. Solids **271,** 157-161 (2000).

[5] S.Q. Man, E.Y.B. Pun and P.S. Chung, Appl. Phys. Lett.**77,** 483 (2000).

[6] R. Balda, J. Fernandez and M. Sanz, Laser spectroscopy of Nd^{3+} ions in GeO_2-PbO-Bi_2O_3 glasses, Phys. Rev. B **61,** 3384 (2000).

[7] Z. Pan, H. Morgan and A. Loper, Infrared to visible upconversion in Er^{3+} doped lead germinate glass: Effect of Er^{3+} ion concentration, J. Appl. Phys. **77,** 4688 (1995).

[8] S.Kohara, H. Ohno, M. Tokata, T. Usuki, H. Morita, K. Suzuya, J. Akola and L. Pusztai, Lead silicate glasses: Binary network-former glasses with large amounts of free volume, Phys. Rev. B **82,** 134209 (2010).

[9] R. Czajka, K Trzebiatowski, W Polewsk, B Koscielsk, S Kaszczyszyn and B. Susla, AFM inestigation of bismuth doped silicate glasses, Vacuum **48,** 213 (1997).

[10] S.R.J. Brueck and X. Long, United State Patent, US6 **436**, 857 (2002).

[11] I. Fanderlik, Optical Properties of Glasses, Elsevier, New York **5**, 92 (1983).

[12] P. Nachimuthu, P. Harikishan and R. Jagannathan, Phys. Chem. Glasses **38**, 59 (1996).

[13] M. Vithal, P. Nachimuthu, T. Banu and R. Jagannathan, Optical and electrical properties of PbO-TiO_2, PbO- TeO_2 and PbO-CdO glass systems, J. Appl. Phys. **81,** 7922 (1997).

[14] J. Tauc, Amorphous and Liquid Semiconductor, Plenum Press, New York, (1974).

[15] B. Carette, M. Ribes and J.I. Souquet, The effects of mixed anions in ionic conductive glasses, Solid State Ionics **9,** 735 (1983).

[16] H. Rawson, Inorganic Glass-Forming Systems, Academic Press, London (1967).

[17] F. Urbach, The Long-Wavelength Edge of Photographic Sensitivity and of the Electronic Absorption of Solids, Phys. Rev. **92,** 1324 (1953).

[18] G. Ambrosone, U. Coscia, S. Ferrero, F. Giorgis, P. Mandracci and C.F. Pirri, Structural and optical properties of hydrogenated amorphous silicon-carbon alloys grown by plasma enhanced chemical vapour de position at various RF powers, Philos. Mag. B **82**, 35 (2002).

[19] V. Dimitrov and S. Sakka, Electronic oxide polarizability and optical basicity of simple oxides, J. Appl. Phys. **79,** 1736 (1996).

[20] C.A. Gressler and J.E. Shelby, Lead fluoroborate glasses, J. Appl. Phys. **64**, 4450 (1988).

[21] T. Banu, K. Koteswara Rao and M. Vithal, Optical, thermal and electrical studies of Nasicon type $Na_2PbZnMP_3O_{12}$ (M=Al,Fe, and Ga) glasses, Phys. Chem. Glasses **44,** 32 (2003).

[22] J. Galy, G. Mennier, S. Andersson and A. Astro¨m, StereoChimie des elements comportant des paires non liei's :Ge(II), As(III), Se(III), Br(V), Sn(II), Sb(III), Te(IV), I(V), Xe(VI), T(I), Pb(II) et Bi(III) (oxides, fluorures etoxy fluorures), J. Solid State Chem. **13,** 142 (1975).

[23] R. J. Gillespie, Molecular Geometry, Van Nostrand Reinhold, London (1972).

[24] J. Duffy, Phys. Chem. Glasses **30,** 1 (1989).

ASYMMETRIC POLY(ETHER-BLOCK-AMIDE)-1657 NANOFILTRATION MEMBRANES FOR PROCESSING OF AQUEOUS SOLUTIONS

B. Venkata Swamy [1], S. Sridhar [2], R.S. Prakasham [3]

[1]Assistant Professor, Department of Biotechnology, B.V.R.I.T, Narsapur, Medak 502313(India)

[2]Principal Scientist, Membrane Separations Group, Chemical Engineering Division, Indian Institute of Chemical Technology, Hyderabad 500007,(India)

[3]Senior Scientist, BEEC Division, Indian Institute of Chemical Technology, Hyderabad 500007,(India)

ABSTRACT

Nanofiltration (NF) is a membrane-based separation process used for treatment of industrial effluents and wastewater recycling. NF membranes have high potential to remove low molecular weight trace contaminants from aqueous and non-aqueous solutions, which cannot be removed efficiently by other conventional treatment methods. In the present study, asymmetric PEBAX-1657 (Poly ether block amide) nanofiltration (FNF) membranes were synthesized by phase-inversion technique from mixed solvent system. These indigenous membranes were characterized by Fourier transform infrared spectroscopy (FTIR), X-ray diffraction (XRD), Thermo gravimetric analysis (TGA) and Scanning electron microscopic (SEM) studies to elucidate the structural, crystallinity, thermal stability, surface and cross-sectional morphologies of the membranes, respectively. The effect of various operating parameters such as permeate flux, total dissolved solids (TDS) and impurity rejection were studied. From the experimental results, an average flux of 68 L/m^2 h (Pebax-2533) and 66 L/m^{2}h (pebax-1657) were obsereved at a constant pressure of 21 kg/cm^2. TDS rejections were found to be 15.18 and 38.56 % respectively. A detailed economic estimation of commercial NF system for feed effluent capacity of 1 m^3/h is presented.

***Keywords:* Characterization, Economic estimation, Flux, Nanofiltration Membrane, Water purification**

I INTRODUCTION

Membrane technology for Nanofiltration (NF) has gained significant importance owing to its lower energy requirements, capital investment, compactness, process safety and environmental viability (Louie et al., 2006). Poly(ether-block-amide) (PEBA) is the general name for a class of thermoplastic elastomers that consist of linear chains of rigid polyamide (PA) blocks and flexible polyether (PE) blocks. This copolymer depending on the PE block concentration becomes hydrophobic or hydrophilic. The hydrophobic grades exhibit high selectivity for extraction of aromatic organic compounds from water by pervaporation (PV) (Nunes et al., 1995). These polymers are rubbery, thermoplastic elastomers with different degrees of hydrophilicity depending on the type of polyamide

(e.g. Nylon- 6, Nylon-12) and polyether (e.g. polyethylene oxide, polytetramethylene oxide) segments (Kim et al., 2001). One way to enhance NF membrane productivity is to reduce the rate of fouling (i.e., the deposition of foreign material) on the membrane surface by coating membranes with a thin, highly water-permeable polyether-polyamide block copolymer (PEBAX) (Chan and Chen, 2004).

Recent studies of polymeric membranes for flue gas applications have focused on improving their performance to allow them to be cost-competitive with solvent absorption (Kujawski and Roszak,2002; Wilks and Rezac, 2002). Structurally asymmetric or composite membranes consisting of a thin separation layer with the support of a micro-porous substrate are used in almost all industrially important gas separation processes (Scholes et al., 2008; Liu, 2004). Lua and Shen (2013) prepared a carbon–silica composite membrane using sol–gel technique, and it surpassed Robeson's upper bound limit for He/N_2, CO_2/N_2 and O_2/N_2 gas pairs. Liu et al. (2005) produced a PEBAX composite membrane to separate CO_2 from N_2 by dip-coating the polymer substrate and investigated the effects of operating conditions on the performance of hollow fiber membranes.

Polyether-block-amide (Pebax) has recently been studied to remove ethanol from aqueous solutions since it provides competitive permselectivity towards specific organic solvents [Scholes etal., 2005; Liu et al., 2009). Pebax is a group of copolymers containing the hard polyamide (PA) segments and soft polyether (PE) segments. In Pebax molecules, PA segments promote mechanical strength while PE segments provide good affinity to organic solvents (Brink et al., 1993; Kou et al., 2003). Depending on the nature and proportion of PA and PE segments, the characteristics of the Pebax polymer can be varied (Gilron et al., 2001; Wang et al., 2006). Generally, the higher proportion of the flexible PE segment, the more organophillic the Pebax polymer is. Among commercially available Pebax polymers (2533, 3533, 4033, 1657, 1740), Pebax 1657 has the highest content of PA segment. It is therefore expected to have the highest hydrophiliicity and maximum separation capacity for the removal of organic compounds from aqueous solutions.

In response to the challenge of developing new membrane materials, PEBAX-1657 NF membrane has been synthesized in this study for removal of inorganic and organic pollutants from industrial effluent and water treatment. The indigenous membrane was characterized by Fourier transform infrared spectroscopy (FTIR), X-ray diffraction (XRD), Thermogravimetric analysis (TGA) and Scanning electron microscopy (SEM) to study the structure, crystallinity, thermal stability, surface and cross-sectional morphology, respectively. The effect of various experimental parameters on water flux and % rejection has been evaluated. The influence of composition in terms of total dissolved solids (TDS), conductivity, turbidity on membrane performance was determined. PEBA membrane was prepared for separation of aqueous mixtures. At first, effect of different parameters on film formation such as ratio of solvents, temperature, composition of coagulation bath (water) and concentration of polymeric solution were studied. The prepared membrane showed good performance for separation of aqueous solutions. A detailed economic estimation of a commercial NF system for processing 1 m^3/h of feed is also presented.

II EXPERIMENTAL

2.1. Materials and methods

Pebax-1657 was chosen to enable greater interaction with H_2O molecules through H-bonding as this grade of the polymer contains 40% amide groups. Pebax-1657 was purchased from Sigma-Aldrich, Mumbai, India. Pure (100%)

ethanol, methanol, citric acid, HCl, EDTA, NaOH and sodium metabisulphite (SMBS) for washing and storage of the membranes were obtained from sd Fine Chemicals Ltd., Mumbai. The indigenous NF membrane of flat sheet configuration having an effective area of 0.015 m^2 was synthesized in the laboratory. Automatic film coater and non-woven fabric support were obtained from Permionics Membranes Pvt. Ltd., Vadodara, India. Colorimeter (Hach-DR-890, Bangalore, India), Conductivity (DCM-900) and pH meters (DPH-504) were procured from Global Electronics, Hyderabad, India, to facilitate analysis of feed, permeate and reject samples.

2.2. Synthesis of Pebax-1657 NF membrane

Solvent resistant membranes of Pebax-1657 were prepared by phase inversion method using a 15 wt.% Pebax-1657 solution by initially adding a small amount of the polymer pellets to a solvent mixture of 90% ethanol and 10% water. After complete dissolution of the initial polymer, remaining polymer was added gradually. The polymer was dissolved at 90 ^oC with rigorous stirring with constant reflux over a period of 6 h. The bubble free polymer dope solution was cast using a doctor's blade on a non-woven polyester fabric support affixed onto a clean glass plate. After 30 seconds the plate was immersed in ice cold water bath to obtain solvent resisiant NF membrane whose total thickness was 103 µm including non-woven fabric thickness. The chemical structure of Pebax-1657 is shown in Fig.1

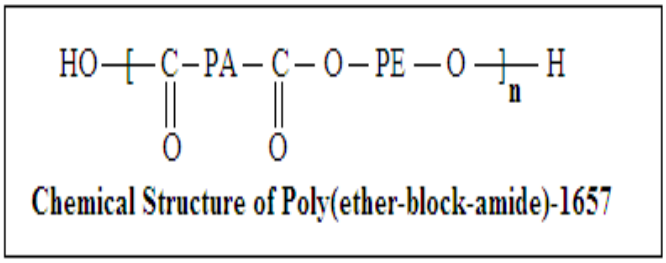


Fig. 1: Chemical Structure of PEBAX-1657

2.3. Description of NF system

The schematic of pilot scale NF system is shown in Fig. 2. A feed tank of 30 L capacity made of stainless steel was provided for storage and supply of the industrial effluent to the system. A polypropylene (PP) prefilter cartridge of 5 µm pore size was installed upstream of the flat sheet membrane module to prevent the entry of suspended solid particles. A high pressure pump (Hironisha, Japan) capable of maintaining a pressure upto 21 kg/cm^2 was used for transporting the feed liquid throughout the system. The pump was run by a 2 HP single phase motor (Crompton, India). The feed tank had a provision for recycling of the concentrate after it passed through a heat exchanger that was installed for maintaining a constant feed temperature (26-28 ^oC). The heat exchanger consisted of a lengthy glass shell, which was circulated with ice cold water while the effluent was flowing in concentric glass coils placed inside the shell which provided large heat transfer area. The reject coming out of the heat exchanger was then recycled to the feed tank as concentrate. A restricting needle valve was provided on the concentrate outlet at a position after the membrane module to pressurize the feed to a desired value which was indicated by a pressure

gauge installed upstream of the needle valve in the reject line. Permeate and concentrate flow rates were measured using rotameters (Venkata Swamy et al., 2013).

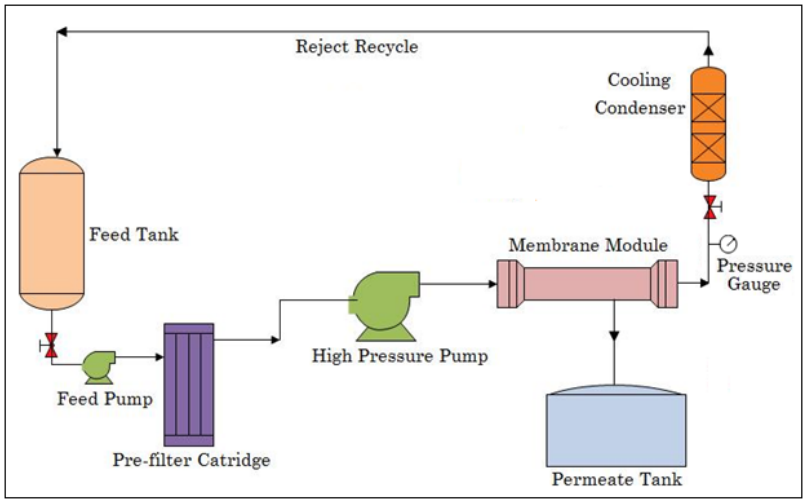


Fig. 2: Process Flow Diagram of NF system

2.3. Experimental procedure

The NF membrane was fixed in the system's flat sheet test cell, cleaned and wetted using distilled water. The experiments were then carried out with deionized water to study the effect of feed pressure on flux. Deionized water was taken in the feed tank and transported to the flat sheet NF membrane module using the high pressure pump. The concentrate outlet was left to flow into a bucket instead of being recycled to the feed tank in order to maintain a constant feed concentration. The system pressure was varied by throttling the needle valve in the concentrate line. The flow rates of permeate and concentrate were measured at each pressure. The stainless steel feed tank was then filled with 30 L of IICT bore well water feed and the system was initially run to remove 2.2 L of distilled water accumulated as dead volume in the system. A sample of initial feed was collected for analysis. The experiment was performed to study the effect of feed concentration on flux and % rejection at an optimum pressure of 21 kg/cm^2 with total concentrate recycle for achieving 65% water recovery. The flow rates of permeate and concentrate were measured at regular time intervals to observe any decline in flux. After a particular water recovery was attained, samples of initial feed, final concentrate and average permeate were analyzed for TDS, conductivity and turbidity values. Finally, the system was cleaned and washed with distilled water to remove solutes from the membrane surface and pores (Venkata Swamy et al., 2013).

2.4. Fouling and its prevention

In general, fouling of membranes is caused by accumulation of suspended solids, salts, microbes and organic materials present in the feed water either on the membrane surface or within the pores (Luo et al., 2012). A solution of citric acid or HCl (1% w/v) was recycled through the system for about 10 min for removal of mineral scales and metal precipitates. A mixture of 1% w/v sodium hydroxide + 0.5% tetra sodium EDTA chelating agent was used to

remove organic scales. On alternate occasions, 0.1% w/v of sodium lauryl sulfate, a surfactant, was added to this alkaline cleaning mixture for polishing the membrane surface**.**

III MEMBRANE CHARACTERIZATION

3.1. Fourier transform infrared (FTIR)

The synthesized NF membrane was characterized for their intermolecular behavior. The membranes were scanned in the range 400–4000 cm^{-1} wavenumber using Nicolet-740, Perkin-Elmer-283B FTIR spectrophotometer (Boston, MA, USA) by KBr pellet method.

3.2. X-Ray diffraction (XRD) analysis

A Siemens D 5000 powder X–ray diffractometer, NJ, USA was used to assess the solid–state morphology of TFC polyamide NF membrane. X-rays of 1.54 Å wavelengths were generated by a CuK–alpha source.

3.3. Thermo gravimetric analysis (TGA)

Thermal stability of NF membrane was examined using a Seiko 220TG/DTA analyzer, Japan in the temperature range of 25–800 °C at a heating rate of 10 °C min^{-1} with continuous flushing using pure N_2 gas flowing at 200 ml/min to determine thermal stability and decomposition characteristics.

3.4. Scanning electron microscopy (SEM)

The surface and cross-sectional morphology of NF membrane was studied by SEM instrument of Model JEOL JSM–6380, LA, USA. In preparing the specimens, the fracture surface and cross-section of the NF membrane, ultraporous substrate and non-woven fabric polyester support were obtained by cutting the membrane samples in liquid N_2 to ensure smooth morphology.

IV ANALYTICAL METHODS

The feed and permeate samples were analyzed for TDS, conductivity and turbidity according to APHA methods (1998). The conductivity of above samples was determined using the digital conductivity meter.

V RESULTS AND DISCUSSION

5.1. Membrane characterization

5.1.1. FTIR

Fig. 3 exhibits the FTIR spectra of NF Pebax 1657 membrane. The prominent peaks of -C=O carbonyl group of stretching vibrations observed in both polyamide (PA) and polyether (PE) groups at 1541cm^{-1}(CO-NH) and 1715 cm^{-1} (-C-O-C-) respectively.

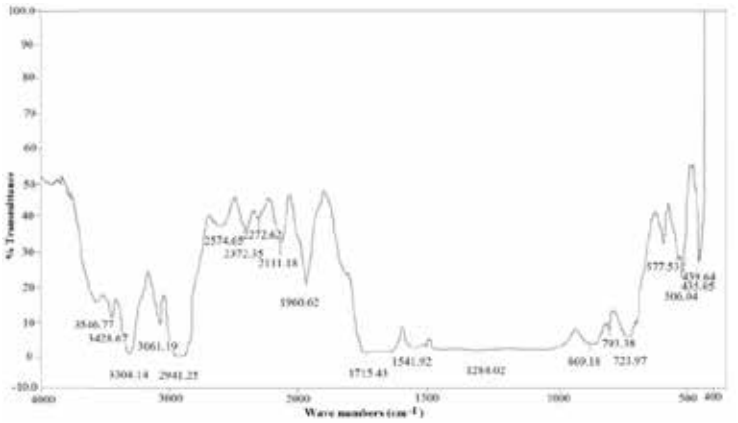


Fig. 3: FTIR Spectra of NF membrane

On other hand bending vibrations of amide group is noted in polyamide at 3304 cm^{-1} and ether linkage of pebax (-C-O-C-) is appears in the polymer chain of 1284 cm^{-1}. These peaks are in accordance with the structure of pebax as confirmed by Kalyani et al (Kalyani et al., 2006).

5.1.2. XRD analysis

X-ray diffraction studies confirms the nature of polymer and and also indicates the intersegmental distance between polymer chains. The X–ray diffractogram of NF pebax membrane shown in Fig. 4 appears to display semi-crystalline nature. XRD pattern shows sharp diffraction peaks at around 17° ,23 and 27° of 2θ scale with *d*-spacing values of 4.92, 3.85 and 3.39 Å, corresponding to the functional groups of the copolymer. The Spectrum support the chemical structure of pebax membrane (Krishna et al., 2012).

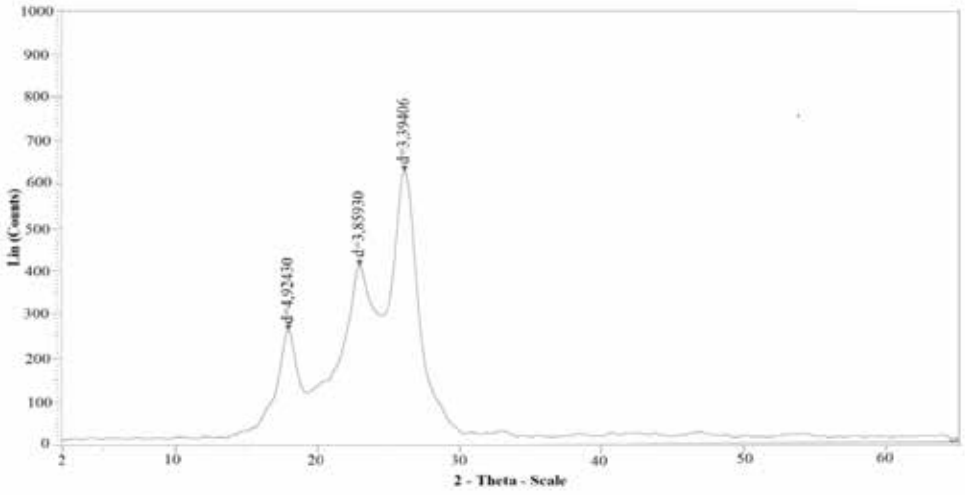


Fig. 4: XRD Spectra of NF membrane

5.1.3. TGA Studies

The TGA curve of NF membrane shown in Fig. 5 indicates the stretching weight loss at 420 °C followed by the decomposition at 470 °C, which is due to the disintegration of molecular chains.

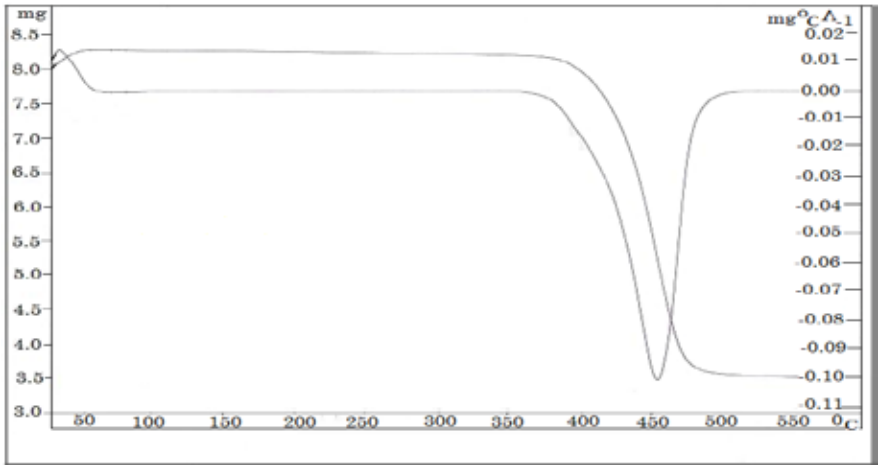


Fig. 5: TGA Curves of NF membrane

From the results it can be concluded that pebax membrane has high thermal stability.

5.1.4. SEM Analysis

The surface and cross-section of NF pebax membrane shown in Fig. 6. Surface morphology of pebax shows nanoporous openings distributed through out the membrane surface. Cross-sectional view of the membrane represents two layers, with the top layer being nanoporous pebax and non-woven polyester fabric support at lower layer with adequate interpretation of nanoporous pebax film.

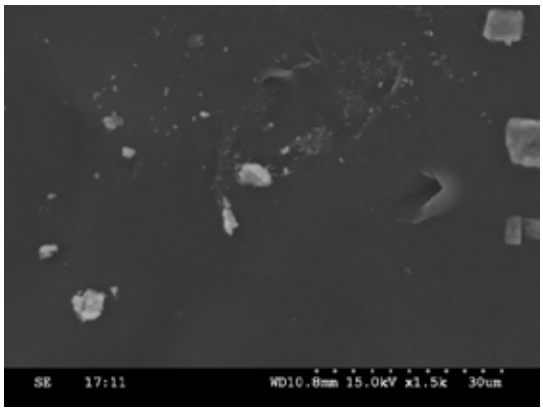


Fig. 6 a: SEM Image of NF Membrane Surface

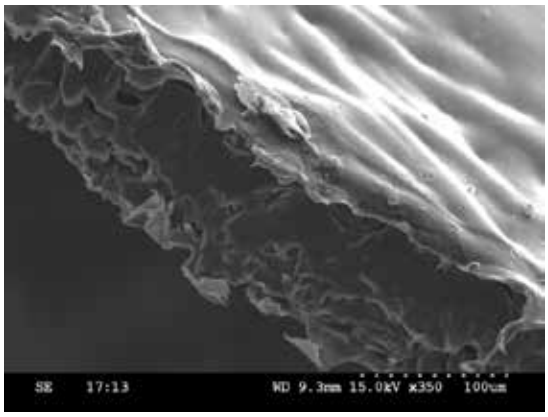


Fig. 6 b: SEM Image of NF Membrane Cross-Section

5.2. Effect of pressure on flux

The effect of pressure on (a) flux and %rejection, (b) conductivity of reject and permeate using NF process is shown in Fig. 7. As expected, a rise in pressure caused an enhancement in both flux and %rejection (Fig. 7a). Since the driving force of the process enhances, it results in enhancement of flux due to increased affinity between H_2O molecules and –CONH moiety of polyamide and –o- functional group of polyether layer.

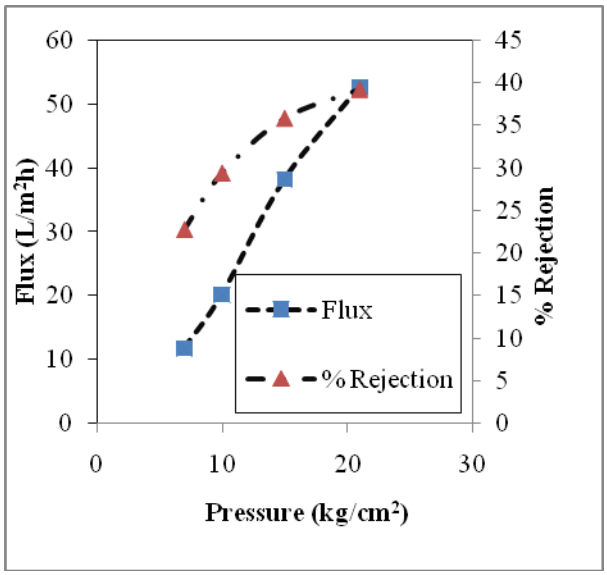


Fig. 7 a: Effect of Pressure on Flux and % Rejection for Bore well Water

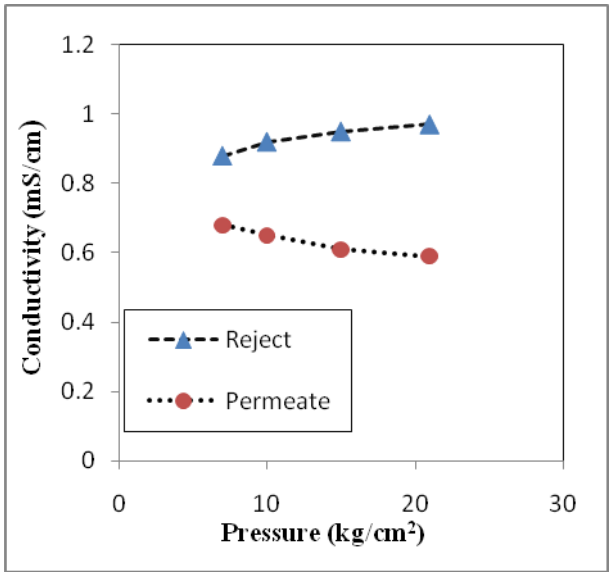


Fig. 7 b: Effect of Pressure on Permeate and Reject Conductivities

The flux was zero at applied pressures less than 3 kg/cm^2 due to high osmotic pressure arising from substantial concentration of dissolved solids in the effluent feed. The conductivity of the permeate reduced with increasing feed pressure (Fig.7b) since the solute flux remains more or less the same irrespective of the applied pressure owing to lack of any interaction with the membrane as described by the solution-diffusion mechanism (Luo et al., 2012). On the other hand H_20 molecules interact with the membrane and get transported at higher pressures.

5.3. Effect of operating time on flux, rejection and conductivity

Effect of time on flux and % rejection of bore water is graphically illustrated in Fig. 8. Flux declined from 65 to 59.61 L/m^{2}h due to increasing concentration polarization near the membrane surface arising from continuous retention of inorganic and organic solutes. On other hand, the rejection decreased from 38.56 to 32.25% over a period of 15 min (Fig. 8). Correspondingly, the conductivity of permeate increased from 0.6 to 0.63 mS/cm and that of reject from 0.89 to 0.93 mS/cm (Fig. 9) at a water recovery of 65% in permeate.

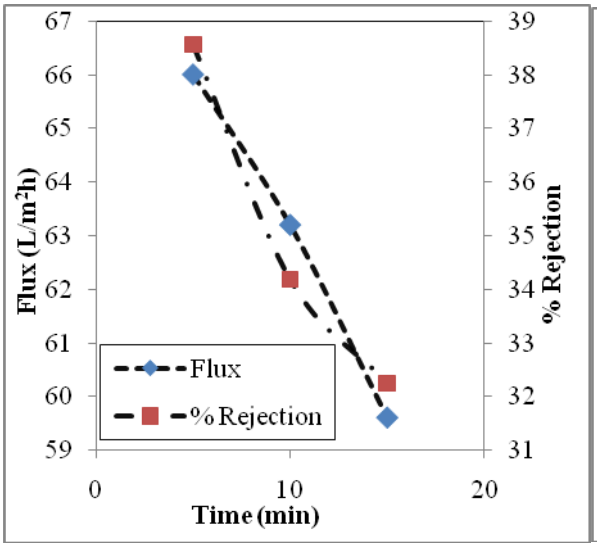


Fig. 8: Effect of operating time on NF Membrane on Flux and % Rejection for bore water

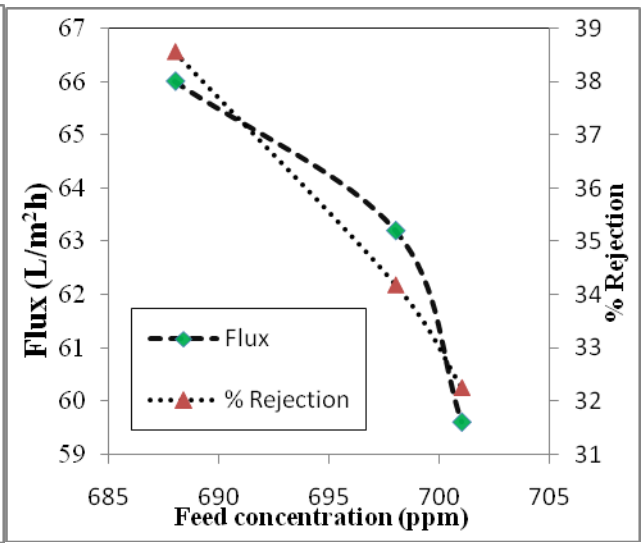


Fig. 9: Effect of Feed Concentration on Flux and Rejection of NF membrane at 21 kg/cm^2

5.4. Separation of TDS, conductivity, turbidity

Table 1 depicts the TDS, conductivity and turbidity values of feed, average permeate and final reject for NF process. A recovery of 65% was obtained at an average flux of 62.60 L/m^{2}h. From experimental observations it can be concluded that the membrane significantly reduces the TDS, conductivity and turbidity present in the feed. Permeate obtained during the process was analyzed as per APHA standards [19] and can be recycled for utilization in agricultural activities ,industrial cooling towers after meeting disposal standards under pollution control board norms.

Table 1: Feed Characteristics of IICT Bore Water and Reduction of TDS, Conductivity, Turbidity and % Recovery

Sample	TDS (ppm)	Conductivity (mS/cm)	Turbidity (FAU)	%Recovery
Feed(Tap water)	603	0.89	100	-
Permeate	576	0.63	6	65.00
Reject	698	0.93	150	-

5.5. Scale-up and economic estimation

5.5.1. Capital investment and list of equipment for NF process

List of equipments and capital costs are provided in Table 2, in which the unit price for major accessories like high pressure centrifugal pump (Grundfos, Denmark), TFC polyamide NF membrane modules with pressure vessel and skid costs are included. The total capital investment for processing of 1 m^3/h of biscuit effluent is approximately 2965 USD (INR 1.77900 Lacs).

Table 2: Capital cost of NF system

Item	Capacity/Size	MOC	Quantity	Total cost (USD)
Membrane housing	--	--	1	200
TFC polyamide membrane module	8” dia × 40” long	--	1	1100
Pressure vessel	35 lpm	--	1	300
Feed pump (2.5 HP)	35 lpm	--	1	150
1 HP high pressure pump	35 lpm	SS	1	800
Skid	17 lpm	SS	1	200
Filter assembly	35 lpm	PP	1	15
Accessories (Rotameters, Valves, TDS/Conductivity, pH meters)	--	--	1 set	200
Total cost				**2965 (INR 1.77900 Lacs)**

(lpm – liter per minute; SS – Stainless steel; MOC – Material of construction; PP – Polypropylene)

5.5.2. Operation and maintenance cost of NF process

Operating and maintenance costs of NF system are given in Table 3, which includes membrane replacement and

Table 3: Operation and maintenance cost of NF system

	1
	0.85
Feed capacity (m^3/hr)	85
Permeate capacity (m^3/hr)	
Module replacement cost	
Number of modules (8’’ dia, 20’’ long)	1
Price per module (USD)	350
Total module replacement cost (USD)	350
Duration of replacement (Years)	3

No. of working hs per day	18
Cost/hr (USD)	0.017
Cartridge replacement cost	
No. of cartridges	1
Price per cartridge (USD)	8
Total cartridge replacement (USD)	8
Duration of replacement (days)	180
No of working hs per day	18
Cost/h (USD)	0.002
Power cost	
Feed pump (kW)	1.12
High pressure pump (kW)	1.85
Dosing system (kW)	0.015
Total power consumption (kW)	2.98
Price per unit (USD) (6 Rs./unit)	0.1
Total power cost (USD)	0.3
Chemical consumption	
Antiscalant dosing (ppm)	5
Dosage (L/h)	0.01
Cost/lit (USD)	6.4
Hourly cost (USD)	0.064
CIP chemicals (EDTA, NaOH, citric acid)	
Frequency (days)	15
Total cost of CIP per hour (USD)	0.045
Labor cost per hr (USD)	0.463
Total operating cost per h (USD)	0.89
Total operating cost per year (USD)	5847.3
Depreciation cost (assuming 10% of capital cost) (USD)	178
Interest (5% of capital cost) (USD)	63
Total cost per year (USD)	6036.3
Permeate	
Quantity (LPH)	850
Operation time (h)	18
Quantity of permeate generated in 1 year (L/yr)	5584500
Cost of permeate per liter (USD)	1.108×10^{-3}
If sold at 4×10^{-3} USD per liter	
Annual Profit (USD)	16301.7

Payback period (yrs)	0.37

prefilter cartridge replacement costs, electric power consumption besides chemicals for cleaning and storage of the membranes. Feed capacity and recovery were assumed to be 1 m^3/h and 65% recovery, respectively. The operating duration was assumed to be 18 h per day and duration of membrane module replacement once every 3 years. Depreciation costs were taken as 10% of total capital investment. The cost per liter of permeate produced was found to be Rs. 0.0846 (1.41×10^{-3} USD) with a payback period of 0.55 years.

VI CONCLUSIONS

The study revealed that indigenously synthesized Pebax-1657 asymetric nanofiltration (NF) membrane can be easily prepared and effectively used for treatment of water. The membrane was characterized by Fourier transform infrared spectroscopy (FTIR), X-ray diffraction studies (XRD), Thermogravimetric analysis (TGA) and Scanning electron microscopy (SEM) to elucidate structural nature, crystallinity, thermal stability, surface and cross sectional morphology, respectively. The flux and % rejection with respect to effluent was evaluated under various operating conditions. The membrane effectively removed total dissolved solids (TDS) and turbidity from feed to a considerable extent. Purification of bore water by NF process is more economical, energy efficient and eco-friendly when compared to reverse osmosis. The robust NF membrane exhibited considerably high water permeability and substantial anti-fouling ability which clearly offers vast scope for treatment of effluents coming from dairy, biochemical, food and bulk drug industrieswhich may contain aggressive solvents that can damage conventional RO/NF membranes.

VII. ACKNOWLEDGEMENTS

We are thankful to Council of Scientific and Industrial Research (CSIR) for granting funds through MATES XII Five Year Plan project to support our research activities and also would like to thank BVRIT, Narsapur, Medak for giving opportunity to the first author for doing his Ph.D. research.

REFERENCES

[1] Louie, J.S., Pinnau, I., Ciobanu, I., Ishida, K.P., Ng, A., Reinhard, M., “Effects of polyether polyamide block copolymer coating on performance and fouling of reverse osmosis membranes”, Journal of Membrane Science, 280, 2006, 762-770.

[2] Nunes, S.P., Sforca, M.L., Peinemann, K.V., “Dense hydrophilic composite membranes for ultrafiltration”, Journal of Membrane Science, 106, 1995, 49-56.

[3] Kim, J.H., Ha, S,Y., Lee, Y.M., “Gas permeation of poly(amide-6-bethylene oxide) copolymer”, Journal of Membrane Science, 190, 2001, 179-193.

[4] Chan, R., Chen, V., “Characterization of protein fouling on membranes: opportunities and challenges”, Journal

of Membrane Science, 242, 2004, 169-188.

[5] Kujawski, W., Roszak, R., "Pervaporative removal of volatile organic compounds from multicomponent aqueous Mixtures", Separation Science and Technology, 37, 2002, 3559-3575.

[6] Wilks, B., Rezac, M.E., "Properties of rubbery polymers for the recovery of hydrogen sulfide from gasification Gases", Journal of Applied Polymer Science, 85, 2002, 2436-2444.

[7] Scholes, C.A., Kentish, S.E., Stevens, G.W., "Carbon dioxide separation through polymeric membrane system for flue gas applications", Recent Patents on Chemical Engineering, 1, 2008, 52–66.

[8] Liu, L., Chakma, A., Feng, X., "Preparation of hollow fiber poly(ether block amide)/polysulfone composite membranes for separation of carbon dioxide from nitrogen", Chemical Engineering Journal, 105, 2004, 43–51.

[9] Lua, A.C., Shen, Y., "Preparation and characterization of polyimide-silica composite membranes and their derived carbon-silica composite membranes for gas separation", Chemical Engineering Journal, 220, 2013, 441–451.

[10] Liu, L., Chakma, A., Feng, X., "CO_2/N_2 separation by poly(ether block amide) thin film hollow fiber composite Membranes", Industrial and Engineering Chemistry Research, 44, 2005, 6874–6882.

[11] Liu, F., Liu, L., Feng, X., "Separation of acetone–butanol–ethanol (ABE) from dilute aqueous solutions by Pervaporation", Separation and Purification Technology, 42, 2005, 273–282.

[12] Gu, J., Shi, X., Bai, Y., Zhang, H., Zhang, L., Huang, H., "Silicalite-filled polyetherblock- amides membranes for recovering ethanol from aqueous solution by pervaporation", Chemical Engineering Journal, 32, 2009, 155–160.

[13] Brink, L.E.S., Elbers, S.J.G., Robbertsen, T., Both, P., "Anti-fouling action of polymers preadsorbed on ultrafiltration and microfiltration membranes", Journal of Membrane Science, 76, 1993, 281-291.

[14] Kou, R.Q., Xu, Z.K., Deng, H.T., Liu, Z.M., Seta, P., Xu, Y.Y., "Surface modification of microporous polypropylene membranes by plasma-induced graft polymerization of {alpha}-allyl glucoside", Langmuir 19, 2003, 6869-6875.

[15] Gilron, J., Belfer, S., Vaisanen, P., Nystrom, M., "Effects of surface modification on antifouling and Performance properties of reverse osmosis membranes", Desalination, 140, 2001, 167-179.

[16] Wang, T., Wang, Y.Q., Su, Y.L., Jiang, Z.Y., "Antifouling ultrafiltration membrane composed of Polyethersulphone and sulfobetaine copolymer", Journal of Membrane Science, 280, 2006, 343-350.

[17] Venkata Swamy, B., Madhumala, M., Prakasham, R.S., Sridhar, S., "Nanofiltration of bulk drug industrial Effluent using indigenously developed functionalized polyamide membrane", Chemical Engineering Journal, 233, 2013, 193–200.

[18] Luo, J., Ding, L., Wan, Y., Paullier, P., Jaffrin, M.Y., "Fouling behavior of dairy wastewater treatment by nanofiltration under shear-enhanced extreme hydraulic conditions", Separation and Purification Technology, 88, 2012, 79–86

[19] Standard methods for examination of water and wastewater, 20th edi, APHA-AWWA-WPCF, USA, 1998.

[20] Kalyani, S., Smitha, B., Sridhar, S., Krishnaiah, A., "Blend membranes of sodium alginate and hydroxyethyl-ethylcellulose for Pervaporation based enrichment of t-butyl alcohol", Carbohydrate Polymers, 64, 2006, 425–

432.

[21] Krishna, A.R., Dev, L., Thankamani, V., "An integrated process for Industrial effluent treatment and Biodiesel production using Microalgae", Research in Biotechnology, 3, 2012, 47–60.

Biographical Notes

Mr. B. Venkata Swamy is working as a Assistant Professor in Biotechnology Department, B.V.R.I.T, Narsapur, Medak and presently pursuing Ph. D from JNTU, Hyderabad, India.

Dr. S. Sridhar working as a Principal Scientist and Project Leader, Membrane Separations Group, Chemical Engineering Division, Indian Institute of Chemical Technology, Hyderabad, India.

Dr. R.S. Prakasham is working as a Senior Scientist, BEEC Division, Indian Institute of Chemical Technology, Hyderabad, India.

AN EFFICIENT STRATEGY OF PREPROCESSING FOR OBTAINING KNOWLEDGE FROM WEB USAGE DATA

Manjula S[1], Rashmi M.J[2] and Varsha.D[3]

[1]*Research Scholar, Dept. of Studies in CS, Solapur University, Solapur (India)*

[2,3] *Student 4th sem MSc Computer Science, Davangere University, Davangere (India)*

ABSTRACT

The World Wide Web (WWW) is a collection of huge amount of Web Usage Data. The process of extracting the relevant data from Web Usage Data is known as Web usage mining. This data must be assembled into a consistent and comprehensive view, in order to be used for further steps of the Web Usage Mining. However, often most of this data are not of much interest to most of the users. Due to this abundance, it became essential for finding ways in extracting relevant data from this ocean of data, hence several researches have been done and researchers proposed an significant and unifying area of research is known as Web mining. As most in data mining technique the data preprocessing involves the removing of irrelevant and inconsistent data, but proper data cannot be achieved without implementing proper preprocess techniques. In this paper we are mainly focusing on the complete preprocessing techniques, such as- data fusion, data cleaning, user identification, session identification, data formatting and summarization. These are the activities used to improve the quality of the data by reducing the quantity of data. This methodology will reduce the size of the data from 75% to 85% from its original data size in Web Usage Mining.

Keywords- Data Preprocessing, Data Formatting and Summarization, Session Identification, User Identification, Web Log Data, Web Usage Mining.

I. INTRODUCTION

In the today's world the use of the Web or Internet is increasing enormously with the huge amount of the data. Details available in the web are related to all most all the fields. The key thing is the user must extract only the details which are necessary for his related work. The process of extracting the details or information from the web is known as *"web mining"*. This web mining is the combination of both data mining and World Wide Web. The web mining is categorized into three areas, such as *web content mining, web structure mining and web usage mining.* The web usage mining is relatively independent while compare with the web content mining and web structure mining, but it's not isolated, this web usage mining technique helps in preprocessing of user data, discover the user's pattern and try to predict the user's behavior figure 1 represents the different steps of

Web Usage Mining. In this paper we are mainly focusing on one of the steps of the web usage mining, that is *preprocessing* technique. The important aim of the preprocessing technique is to remove inconsistent data, redundant data, noisy data where all these data will be present in the web server log files along with useful information, by applying preprocessing techniques we can reducethe size of the web server log file from 75% to 85% of the original size of the web log file, finally the output of preprocessing that is interesting patterns can be transformed to a relational database models.

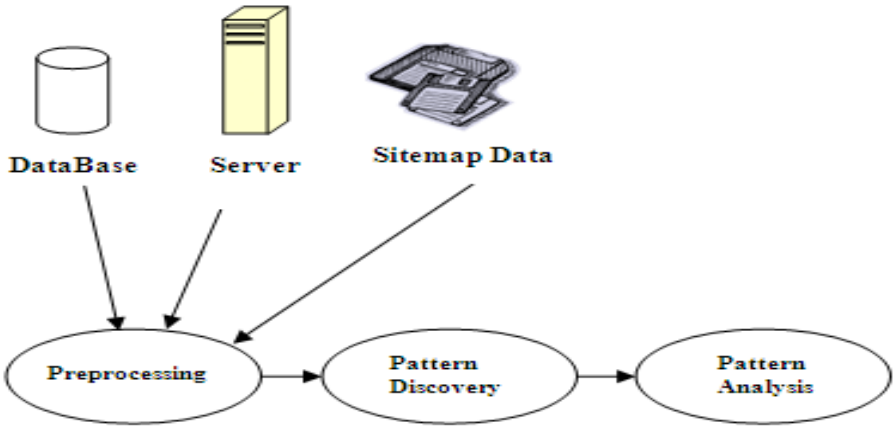


Figure1: General Web Usage Mining Process

II. LITERATURE SURVEY

The data preprocessing is a set of operations that process the available sources of information and lead to the creation of an ad-hoc formatted dataset to be used for knowledge discovery through the application of mining techniques [8]. The main purpose of preprocessing is to produce result that can be used to improve and optimize the content of a site [24].This mining technique will reduces the quantity of data by increasing the quality of data, this will reduce nearly 80% of the original size of the log data[18]. The major steps involved in the preprocessing are data merging, data cleaning, user identification, session identification and data formatting and summarization. The data merging is the technique that combines web log data from multiple sources to retrieve additional information with respect to identification of users and sessions [15].Several papers have different explanation which gives the depth knowledge on this data merging techniques. The data cleaning refers to the removing of irrelevant information which is useless for mining purposes from the HTTP server log file[19]. User identification, is to identify the visitors who access web site and which pages are accessed, it is necessary to distinguish among different users, since user may visit a site more than once [14]. The goal of session identification is to divide the each uses who visit the web site more than once into individual sessions, this can be achieved by using the technique of timeout to break a user's click-steam into session [6]. After the identification of session obtained preprocessed data need to be stored in the relational database module which can be used to properly format the session or transaction for the type of data mining to be accomplished[16].The Data summarization concerns with computation of aggregated variables at different abstraction level[4]. In this

paper we are mainly focusing on different treads of preprocessing by providing appropriate algorithm for implementing preprocessing strategy for obtaining knowledge from Web Usage Data.

III. PREPROCESSING

The data preprocessing technique is the set of operations that extract the information from the available source of information which can be used for further steps of Web Usage Mining. The available Web Usage Data is usually consists of irrelevant data, inconsistent data, noise, etc. For example, the user requests for graphical page content and requests for any other file which might be included into a web page or even navigation sessions performed by robots or web spiders which need to be removed, hence this preprocessing strategy have been proposed to extract the required data from the Web Usage Data. Data preprocessing is predominantly significant phase in Web Usage Mining due to characteristics of web data and its association to other related data collected from the multiple source. The data provided by the data source can be used to construct a data model, the web server is the richest source of data, web serverwill collect large amountof data from web sites and these data is stored in the web log files.The web log files are act as input for this preprocessing. These log files are available in web servers.These log files contains the multi-user details in the log file format. The two main log file formats are Common Log Format (CLF) and Extended Common Log Format (ECLF).A typical line in CLF is shown in figure 2.Recently W3C [W3C log] has represented an improved format for Web server log known as Extended Common Log Format (ECLF), this ECLF log file consists of two more fields then the CLF, the referrer (the URL client was visiting before requesting the URL) and user agent (the software that claims to be using) fields. A typical line in ECLF is shown figure3. Both the CLF and ECLF consist of following fields: *IP address:* This represents the address of the client's host name. *Rfc:* It is the remote login name of the user but most of the time it's not available because systems are usually identified by IP address, if remote login name is available then it can be used (here it's not available hence minus sign has been mentioned). *Authorized User:* This is available only when the WWW documents or pages are password protective or if any web site has restricted for the public use then the person who wish to access this sites then the authorization name and password will be provided for those visitors. *Date and time:* This provides the details of the site visitors on what date & time visitor access this sites. *Request:* The details of the website which browser has requested from the website. *Status:* The request status code will be provided so that it will identify whether the requested page or document is ok to access or its under development, this can be achieved by representing code numberin the log file format (Status codes uses in the log files are: 200 OK, 201 Created, 202 Accepted, 400 Bad request, 404 Not found, 403 Forbidden, 204 No Content, 401 Unauthorized, 304 Not modified, 301 Moved Permanently etc…,). *Byte:* Requested web page size in the bytes format. *Referrer:* This represents which is the URL of the Web page containing hyperlink that will display the current document. *User agent:* This will help to find the browser and operating systems which is used for this request to be displayed.These web log files will be the input for the preprocessing strategy; Preprocessing involves Data fusion, Data cleaning, User identification, Session identification, Data Formatting and Summarization.

The goal of preprocessing is to transform the raw data contained in database of web log file into a transaction log file. The data preprocessing presents a number of unique challenges which lead to a variety of algorithm and heuristic techniques for preprocessing tasks. Figure 4represents the different steps of preprocessing.

<ip_addr><base_url><date>
<method><file><protocol>
<code><bytes>

Figure 2: Common Log Format (CLF)

<ip_addr><base_url><date><method>
<file><protocol><code><bytes>
<referrer>

Figure 3: Extended Common Log Format (ECLF)

3.1 Data fusion

Data fusion refers to the use of techniques that combines data from multiple sources to retrieve additional information with respect to identification of user and session, than if they are retrieved from a single data source.

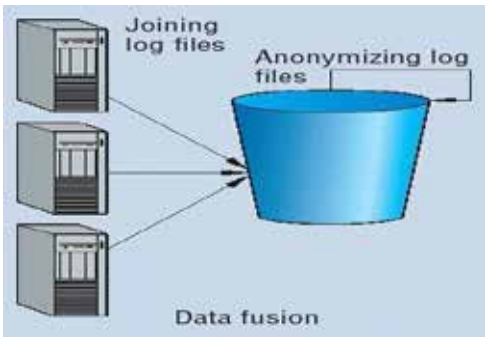


Figure5: Data fusion representation

The different log files are put together and forms a single log file. There is a slight improvisation technique. In this process as it is not only a combining of log files, at the same time it will sort all the entries with respect to time stamp, so it reduces the time and resources to apply and implement separate steps for the sorting of this huge of fusion data of different log files. Figure 5 represents the data fusion method. The fusion and sorting problem is formulated as "The set of log files are provided that is considered as L_f= {L_1, L_2...L_n} are combined into single log file L". The data variable required to store the values or an array [12]. The data fusion technique can be implemented with the help of algorithm given in figure 6.

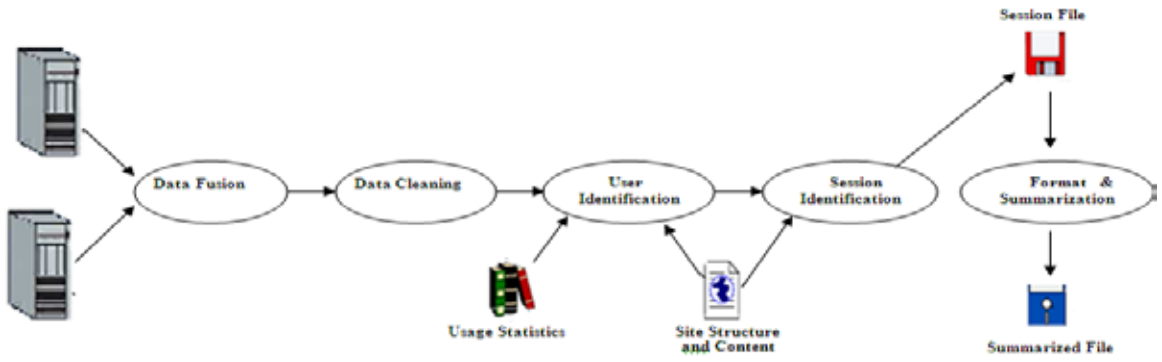


Figure 4: General steps for Preprocessing

Algorithm for Data fusion with different web server log at Internet Service Provider (ISP)

```
//Input: different log file generated by proxy servers.
//Output: single file containing combined output in timestamp order.
For (each log file input) do
        Read log file one by one; do until end of the file reached.
        Read record one by one from current log and put into output file.
        Increment the pointer to point the next record in current log file.
    Increment the pointer in output file to store the next record.
End of for loop.
Sort the output log file entries in ascending order by access time.
Return output.
```

Figure 6: Data fusion algorithm

3.2 Data cleaning

The process of filtering or removing irrelevant item, noise, inconsistent data which is stored in the log files that may not be useful for further web usage mining process, is known as data cleaning. The object of this is to obtain only usage data file request that the user did not explicitly request can be eliminated, this technique also removes the spider navigations requests of web robots. The log entry files such as *.gif, *.jpeg, *.?, *.GIF, *.JPEG, etc. can be removed [19]. Invalid requests from the proxy log files that refer to either internal server error with code and server sides, errors are also removed from the proxy server log files. This technique can be implemented by using the algorithm of data cleaning represented in figure 7. Consider an example for data cleaning, the table I containing all the details of the log file with different extension, by using data cleaning algorithm we need to remove all the irrelevant data from this log file. If the status code of the requested page is greater than 299 or less than 200 then those records need to be deleted, or if the requested method is not in {GET, POST} status then those records need to be removed from the log files. The outcome obtained from this is presented in table II.

Algorithm for Data Cleaning.

```
//Input: Raw log file generated by the proxy server.
//Output: Preprocessed log file with relevant entries consists of tuple (timestamp, IP, URL).
//Constraint: Log file stored in text file must transform into database for further processing.
Read record in database for each record in database.
Read files URL field in proxy server log; the requested objects is the URL field.
If requested URL field contains or end with substring {*.jpeg, *.gif, *.css, *.?}
        Then remove the records.
Else if response code is >299 or < 200
        Then remove records.
Else if requested method not in {GET, POST} status
        Then remove records.
Save record in output.
End of if.
```

Figure 7: Algorithm for Data Cleaning

No	Object Type	Unique Users	Requests	Bytes In	% of Total Bytes In
1	*.gif	1	46	89.00 KB	0.50%
2	*.js	1	37	753.95 KB	4.40%
3	*.aspx	1	34	397.05 KB	2.30%
4	*.png	1	31	137.67 KB	0.80%
5	*.jpg	1	20	224.72 KB	1.30%
6	Unknown	1	15	15.60 KB	0.10%
7	*.ashx	1	15	104.79 KB	0.60%
8	*.axd	1	13	274.81 KB	1.60%
9	*.css	1	8	71.78 KB	0.40%
10	*.dll	1	7	26.41 KB	0.20%
11	*.asp	1	4	1.26 KB	0.00%
12	*.html	1	3	2.17 KB	0.00%
13	*.htm	1	2	69.87 KB	0.40%
14	*.pli	1	2	24.92 KB	0.10%

Table I: Web Log File with different extension before implementing data cleaning algorithm.

No	Object Type	Unique Users	Requests	Bytes In	% of Total Bytes In
2	*.js	1	37	753.95 KB	4.40%
3	*.aspx	1	34	397.05 KB	2.30%
4	*.png	1	31	137.67 KB	0.80%
5	*.jpg	1	20	224.72 KB	1.30%
6	Unknown	1	15	15.60 KB	0.10%
7	*.ashx	1	15	104.79 KB	0.60%
8	*.axd	1	13	274.81 KB	1.60%
10	*.dll	1	7	26.41 KB	0.20%
11	*.asp	1	4	1.26 KB	0.00%
12	*.html	1	3	2.17 KB	0.00%
13	*.htm	1	2	69.87 KB	0.40 %
14	*.pli	1	2	24.92 KB	0.10%

Table II: Web Log File with different extension after implementing data cleaning algorithm.

3.3 User identification

The main task of user identification is to identify the user who access web site and which pages are accessed. This deals with associating page reference with different users; this reduces network traffic and improves performance. In most of the cases web log file provides only the computer IP address and user agent (if web log is using ECLF), then users are identified with the IP address. Some web site requires user registration for the user identification purpose, but due to privacy reason many users prefer not to browse those sites which require registration and logins [19]. Hence in most of the cases user is identified by the IP address of the computer, but different users can use the same IP address for browsing purpose, in this case we can distinguish the user by the software tool or by the operating system used by the user.

3.4 Session identification

The goal of session identification is to divide thepage accesses of user into individual sessions; at present the methods to identify user sessions include timeout mechanism and maximal forward reference. A user session is a directed list of page accesses performed by an individual user during the visit in a web site [4]. For the identification of user session the timeout has been fixed, if the time between page requests exceeds a giving time limit then it is assumed that the user is starting new session. A user may have a single (or multiple) session(s) during a period of time. Thus the session identification problem is formulated as "Given the Web log file *Log*, capture the web users navigation trends, typically expressed in the form of Web users sessions". In this paper we have provided the "timeout threshold" to define the user's sessions. Figure 8 represents Session identification algorithm.

```
Session_Gen ()
  { Let Hi={f1,f2,..., fn} denote the time-ordered session
    history
    Let lj, fj, rj, and tj denote log entry, request, referrer, and
    time (at which the request was received) respectively.
    Let τ denote the session time-out (Usually 30 minutes)
    Sort the Log data by IP address, agent, and time.
    for each unique IP/agent combinations do {
      for each lj do {
        if ( ( (tj-tj-1) > τ ) or ( rj ∉ {H0,H1,..., Hm}) )
        then increment i, add lj to Hi
        else  assign = Distance (H, rj), add lj to Hassign
      }
  }}
```

```
Distance (H,f)
{Let H, denote the time ordered session history.
      Let f denote the page file
      Set min = ∞
      For each Hi∈H do
            If(f ∈ Hi)
            Dj = Hi. size() - Hi.index(f)
            tj = Hi. tn – Hi.tf
            if(dj<min) then
                  assign = i,
                  min = dj
            else
                  if(dj = min) then
                  if(tj< tassign) then
                        assign = i
      return assign
}
```

Figure 8: Algorithm for Session Identification Figure 9: Algorithm for Distance function

The session identification algorithm checks to see the session time-out or if the referring file is not present in any of the open session histories. If so, a new session is opened. Since the Log has already been stored by IP address/Agent, all of the open session histories are potential candidates for the page file access being processed. 'Session_Gen' algorithm calls the 'Distance' function finds the history that most recently accessed f. The 'Distance' function shown in figure 9 given a list of histories and page file f. Finally, if the times are equal, by default, the access is assigned to the history with the lower index. A random assignment could be used to break ties. The index of the history that f should be assigned to is returned by the 'Distance' function.

3.5 Data Formatting and summarization:

It is the terminal step of preprocessing technique,the preprocessed log entries are represented using the final preparation module. This can be used to properly format the obtained preprocessed session or transaction file for the different instance of the web usage mining techniques.The data generalization method is applied at the request level and aggregated for visits and user sessions to completely fill in the database [20]. The data summarization concerns with the computation of aggregated variables at different abstraction levels; these aggregated variables are later used in the data mining step. Table III represents the sample database in session table. This formatted database table is very important for further steps like clustering, pattern discovery, by using this table one can easily come to know about user's session details with IP address and URL accessed by this we can generate the matrix representation for the clustering purpose.

Session Id	IP Address	Date &Time	URL Accessed
1	120.33.med.umich.edu	1995-08-03 22:42:31	/history/apollo/apollo-13/apollo-13.html
1	120.33.med.umich.edu	1995-08-03 22:43:03	/facilities/lc39a.html
1	120.33.med.umich.edu	1995-08-03 22:43:42	/facilities/mlp.html
1	120.33.med.umich.edu	1995-08-03 22:45:01	/facilities/tour.html
2	128.101.144.178	1995-08-03 23:15:10	/shuttle/missions/sts-69/mission-sts-69.html
3	128.102.143.201	1995-08-04 03:15:22	/shuttle/missions/sts-64/mission-sts-64.html
4	128.102.143.212	1995-08-04 03:11:51	/shuttle/missions/sts-69/mission-sts-69.html
4	128.102.143.212	1995-08-04 03:13:15	/facilities/vab.html
4	128.102.143.212	1995-08-04 03:14:28	/shuttle/missions/sts-71/mission-sts-71.html
4	128.102.143.212	1995-08-04 03:15:11	/shuttle/missions/missions.html
4	128.102.143.212	1995-08-04 03:15:18	/shuttle/missions/sts-70/mission-sts-70.html
4	128.102.143.212	1995-08-04 03:16:35	/shuttle/missions/sts-72/mission-sts-72.html
4	128.102.143.212	1995-08-04 03:17:10	/shuttle/missions/sts-72/sts-72-info.html
4	128.102.143.212	1995-08-04 03:17:24	/ksc.html
5	128.102.202.133	1995-08-03 23:13:02	/shuttle/missions/missions.html
5	128.102.202.133	1995-08-03 23:15:30	/shuttle/missions/missions.html
5	128.102.202.133	1995-08-03 23:15:51	/shuttle/missions/missions.html
5	128.102.202.133	1995-08-03 23:16:02	/shuttle/missions/missions.html
6	128.102.236.36	1995-08-03 23:16:45	/history/apollo/apollo-13/apollo-13.html
7	128.102.86.216	1995-08-04 02:23:24	/shuttle/missions/sts-70/mission-sts-70.html

Table III: Sample database in session

IV. IMPLEMENTATION AND RESULTS

Preprocessing methodology has been implemented on the web log files which is downloaded from the NASA website, this is the input for our experiment, these log files which consist of all the irrelevant data, inconsistent data, web spiders, robotic movements which always invoke automatically at the time of browsing hence preprocessing techniques is implemented on these log files which reduces the size of the log file from 75% to 85% of its original size of web log file. We have used JAVA for the implementation of preprocessing methodology. The main part of the implementation is file control which is represented in figure 10; this code segment represents the access control to perform the preprocessing strategy. Before implementing preprocessing, first we need to check whether the requested file name is present in the log files, this code helps to check the requested file exists or not, if it exists then it will allow to read the file, otherwise it will reflects the

error message. Figure 11 represents code segment for data fusion, it will check the size of the log file and it will add the items (if multiple log files are available) for further steps of the preprocessing. After merging the log file data cleaning method will be implemented, here it checks the status of the log file, if file contains the inconsistent data then it will be removed and cleaned data will be stored in the array. Figure 12 represents code segment for data cleaning. Figure 13 represents session identification code segment, in this "TimeOut" is fixed, based on the timeout session will be generated. Figure 14 represents the snapshot of the results which we obtained after the implementation. Table IV represents the result after the implementation of preprocessing methodology; here the size of the log file is reduced from 75% to 85% of its original size. The preprocessed data which is obtained from this implementation can be further used for the pattern discovery purpose in Web Usage Mining.

```
Public static boolean checkFileName(String fName)
{File fileName= new File(fName);
 If(fileName.exists())
 {          if(fileName.isFile())
            {          if(fileName.canRead())
                                  return true;
                       else
                       JOptionPane.showMessageDialog(null,"WUM:Error!");
            }else
            JOptionPane.showMessageDialog(null,"WUM:Error!Specify location");
 }else
 JOptionPane.showMessageDialog(null,"WUM:Error!Specify location");
 return(false);
}
```

Figure 10: JAVA code for File Control

```
Public int addItem(String item)
{          int size;
           If(items.size()=0 || !tems.contains((String)item))
           {          size=itme.size();
                      ItemRef newItemRef = new ItemRef();
                      newItemRef = new ItemRef();
                      newItemRef.setItemID(size+"");
                      newItemRef.setItemName(item);
                      itemsRef.add(newItemRef);
                      items.add((String)item);
           }else
           Size = item.indexOf((String)item);
           return size;
`
```

Figure 11: JAVA code for Data Fusion

```
Public LinkedList removeInfrequentItemsets(LinkedList itemsFreq.int minSup)
{       ListIterator itItemsFreq = itemsFreq.listIterator(0);
        ItemSet newItemsetElements;
        LinkedList indexesToDelete = new LinkedList();
        While(itItemsFreq.hasNext())
        {       newItemsetElements = (ItemSet)itItemsFreq.next();
                If(newItemsetElements.getSupport() < minSup)
        IndexToDelete.add((newInteger(itemsFreq.indexOf(newItem        setElements))))
        )for(int count = indexes To Delete.size()-1; count >=0; count--)
        {       itemsFreq.remove(((Integer)indexes ToDelete.get(count).intValue());
        }return itemsFreq;
}
```

Figure 12: JAVA code for Data Cleaning

```
Public int accessLogProcessor(JTextArea pResultArea)
{int k=statValues.accessData.size();
 While(k>0)
 {      temp = (AccessLogData)statValues.accessData.get(k-1);
        /* fist check whether the IP address is same or not
         * if yes, check whether the user agent and referer is matched.
         * if yes, add this to the session
         * else look for other access records in queue within the session TimeOut*/
  If(temp.ip.equals(result[0]) && (!temp.identifySession(x, session TimeOut))
&&(temp.compareAgentReferer(result[11], result[10])))
                        { temp.addURL(result[6]); ////
                            /* replace the previous datetime*/
                          Temp.setDateTime(x);
                          Flag = true;
                          Break;
                        } k--;
        } /* end of while*/
   }
```

Figure 13: JAVA code for Session Identification

Log File Format	Total number of records	Interesting records	No. of Session Identified	Original log file size	Preprocessed log file size

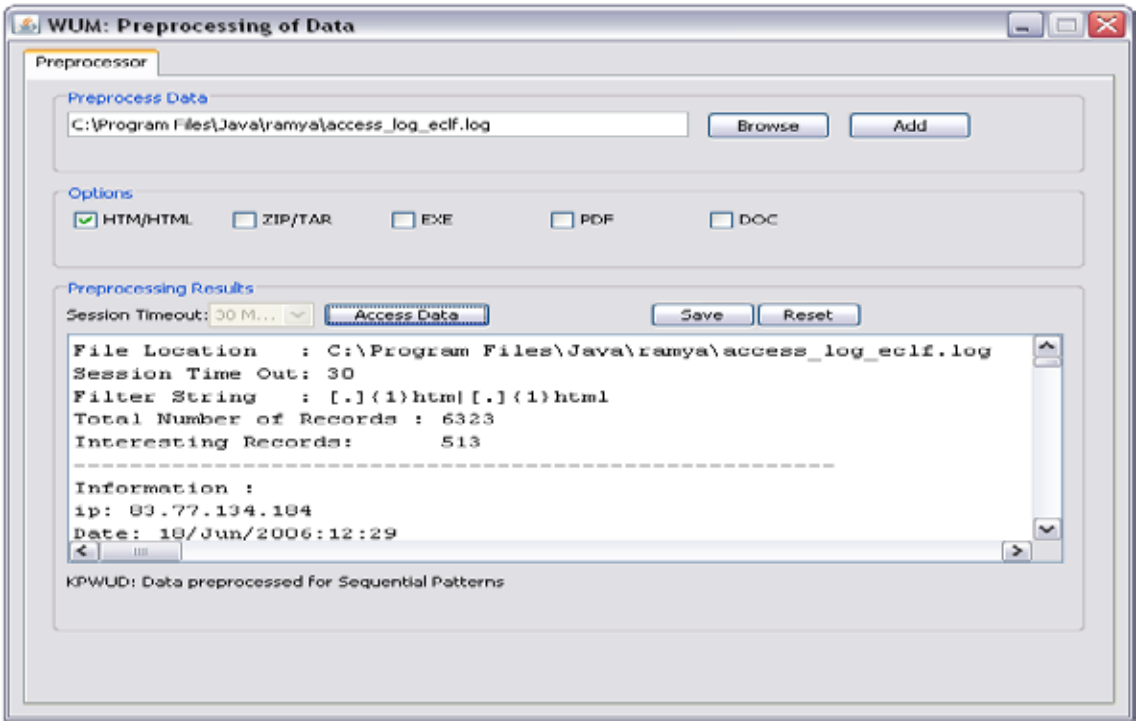
V. CONCLUSIONS

The data collection from different log file will contain all the details about the pages visited by the user during his interaction with the web, along with this log files also contain all the inconsistent, noisy, irrelevant data, hence it is necessary that these log files need to be preprocessed in the précised manner. For the implementation purpose we have downloaded NASA web log file. The preprocessing methodology, such as - data fusion, data cleaning, user identification, session identification, data formatting and summarization steps are implemented on that to obtain the interesting patterns, Table IV represents the result after the preprocessing of data. By this we have proved that this implementation technique reduces the size of the log file from 75% to 85% of its original size, hence it is very efficient strategy of preprocessing for obtaining knowledge from Web Usage Data.

VI. FUTURE WORK

The future work involves getting the solution for the various issues which usually raises at the time of data collection tasks, data transformation tasks and user identification. This also involves various data transformation tasks that are likely to influence the quality of the preprocessed data resulting from the different strategy of preprocessed, this preprocessed data can be used for the pattern discovery which can be further used for various web usage application which as site improvements, business intelligence and recommendation.

Figure 14: Snapshot showing Interesting web patterns after preprocessing.



CLF	5166	2224	1218	419 KB	109 KB
ECLF	6323	513	358	1668 KB	55 KB

Table IV: Results after preprocessing obtained from NASA Log File

REFERENCES

[1]. Configuration files of W3C http. http:/www.w3.org/Daemou/User/Config/ (1995).

[2].W3C Extended Log File Format, http:/WWW.W3.org/TR/WD-logfile.html (1996).

[3]. Yan Wang: Web Mining and Knowledge Discovery of Usage Patterns, CS 748T Project (Part I). Feb-2000.

[4]. Ramya C, Dr. Shreedhara K. S and Kavitha G: Preprocessing: A Prerequisite for Discovering Patterns in Web Usage Mining, 2011.

[5]. J. Srivastava, R. Cooley, M. Deshpande, P-N. Tan: Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data, SIGKDD Explorations, Jan-2000.

[6]. Hui Yu, Zhongim Lu: Analysis of Web Usage Mining.

[7]. Pronalazenje Skrivenog Znanja, Bojan Furla: Data Preprocessing.

[8]. Mathias Gery, Hatem Haddad: Evaluation of Web Usage Mining Approaches for User's Next Request Prediction, Nov-2003.

[9]. Haugua Dai, Bamshad Mobasher: Integrating Semantic Knowledge with Web Usage Mining for Personalization.

[10].Li Chaofeng: Research and Development of Data Preprocessing in Web Usage Mining.

[11]. Aniket Dash: Web Usage Mining: An Implementation.

[12]. Mr. Sanjay Bapu Thakare, Prof. Sangram.Z.Gawali: A Effective and Complete Preprocessing for Web Usage Mining. Vol. 02, No. 03, 2010, 848-851.

[13]. V. Sathiyamoorthi, Dr. V. Murali Bhaskaran: Data Preparation Techniques for Web Usage Mining in World Wide Web- An Approach. Vol.2, No. 4, Nov-2009.

[14]. V.V.R Maheswara Rao and Dr. V. Valli Kumari: An Enhanced Preprocessing Research Framework For Web Log Data Using A Learning Algorithm.DOI:10:5121/csif2011.1101.

[15]. C.P Sumathi, R.Padmaja Valli, and T. Santhanam: An Overview of Preprocessing of Web Log Files for Web Usage Mining. ISSN 1992-8645, E-ISSN 1817-3195.

[16]. R. Cooley, B. Mobasher and J. Srivastava: Data Preparation for Mining World Wide Web Browsing Patterns HER-9554517.

[17]. V. Chitraa, Dr. Antony Selvadoss Thanamani, A Novel Technique for Session Identification in Web Usage Mining Preprocessing. Vol. 34, No.09, Nov—2011.

[18]. Navin Kumar Tyage, Ak. Solonki and Manoj Wadhwa: Analysis of Server Log by Web Usage Mining for Website Improvement. Vol. 07, Issue4, No. 08, July-2010.

[19]. V. Chitraa, Dr. Antony Selvdoss Dawamani: A Survey on Preprocessing Method for Web Usage Data Vol. 07, No.03, 2010.

[20]. Zhiqiang Zheng, Balaji Padmanabhan, Steven O. Kimbrough: On the Existence and Significance of Data Preprocessing Biases in Web Usage Mining. Vol. 15, No. 02, Spring 2003.

[21]. V. Sathiyamoorthi and Dr. Murali Bhaskaran: Data Preprocessing Techniques for Pre-Fetching and Caching of Web Data through Proxy Server. Vol. 11, No. 11, Nov-2011.

[22]. Lukas Cenovsky: web Usage Mining on is.muni.c2 Master's Thesis.

[23]. Suneetha K.R, Dr. R. Krishnamoorthi: Data Preprocessing and Easy Access Retrieval of Data through Data ware House. Oct-2009, ISBN: 978-988-17012-6-8.

[24]Mohd Helmy Abd Wahab, Mohd Nozali Haji Mohd, Hafizul Fahri Hanafi, Mohamad Farshan, Mohamad Mohsin: Data Preprocessing of Web Server Logs for Generalized Association Rules Mining Algorithm-2005.

[25]. Jiawei Han and Micheline Kamber: Data Mining Concepts and Techniques. ISBN-81-8147-049-4

IMAGE CODING, PACKETIZATION AND CHANNEL CODING FOR COMMUNICATION OVER WIRELESS NETWORKS – A REVIEW

Bharathi Gururaj[1], G. Sadashivappa[2]

[1] *Assistant Professor, Department of Electronics and Communication Engineering, ACS College of Engineering, Mysore Road, Bangalore, (India)*

[2] *Professor, Department of Telecommunication Engineering, RV College of Engineering, Mysore Road, Bangalore, (India)*

ABSTRACT

This paper is an exhaustive literature survey on the various wireless networking standards available and data transmission mechanisms currently employed. We discuss data representation, wireless network standards, image transmission mechanisms and two main channel encoding algorithms – Turbo codes and LDPC codes.

Keywords– Packetization, PSNR, QOS, SPIHT, Wireless Networks

I. INTRODUCTION

With the introduction of 3G wireless communication systems, together with phenomenal growth and popularity of internet, wireless multimedia communication is predicted to grow rapidly in near future. The transmission of multimedia data over wireless channels has become increasingly important during the last decade, especially after Wireless LAN networking has become a global standard for high data rate, low mobility users. Third generation cellular systems, such as IMT – 2000 and UMTS are becoming an important endpoint to provide multimedia data to final users, even at lower data rate. However, representing multimedia data requires a large amount of information, leading to high bandwidth, computation energy consumed in processing information to be transmitted and communication energy consumed when in wirelessly transmitting information.

The large requirements for bandwidth and energy consumption are significant bottlenecks to wireless multimedia communication. Multimedia system incorporates continuous media like voice, video, and images. This implies the need for multimedia systems to handle data with strict timing requirements and at high rate. This multimedia content provides rich information to consumers, but also with information it poses challenging problems of management, delivery, access and retrieval because of its data size and complexity. Most of these representations contain large amounts of redundancy can exist in various forms. It may exist in various form of correlation spatially close pixels in an image are generally also close in value.

Low bandwidth and limited channels are one of the most challenging issues that every nation is facing. Since signals and data to be sent via a channel are available in huge amounts and limited spectrum is available data needs to be compressed at the transmitter and expand at receiver side and sometimes even lead to false reconstruction at the receiver side. This compression and expansion leads to distortion of signals. Thus the performance of the system decreases and hinders in smooth operation of the system. If the channel is wireless

the situation becomes even worse. Different amount of noise enter the channel and corrupt data. If a digital image is sent via wireless channel it has to be compressed at transmitter because of limited bandwidth. At the receiver it may get distorted due to noise and during image expansion it might deviate from the original form and data may be lost while traversing. The key to compressing data is to the distinction between the data and information. In this wireless multimedia framework, the streaming of images requires some care on protection of compressed frames because random bit errors or packet losses introduced by the channel may corrupt critical information for decoding of images up to a point that no data at all can be decoded. Channel coding at code stream level plays a major role in preventing the transmission channel from introducing unrecoverable errors. Forward Error Correction is one of the possible solutions when there is no return channel is available for retransmission of packets.

A variety of error-resilient techniques for image transmission have been recently proposed in literature. Most are based on the state of art SPIHT source coder which generates embedded bit streams in which lower rates are prefixes of higher rates.

II. DATA REPRESENTATION

An image is a positive function on a plane. An image may be defined as a two dimensional function f(x, y), where x and y are spatial or plane coordinates, and amplitude of f at any pair of coordinates (x, y) is called the Intensity or Gray level of image at that point[20].

When x, y, and intensity values of f are all finite discrete quantities, we call the image as digital image. A digital image f(x, y) described in a 2D discrete space is derived from an analog image f(m, n) in a 2D continuous space through sampling process that is frequently referred to as digitization. Digitizing the amplitude values is called quantization. Digital image is composed of a finite number of elements each of which has a particular location and value. These elements are called picture elements, image elements, pels and pixels. Pixel is the used most widely to denote the elements of a digital image.

III. REPRESENTATION OF DIGITAL IMAGES

Let f (m, n) represent a continuous image function of two continuous variable, m and n. We convert this function into a digital image by sampling and quantization. We sample the continuous image into a 2D array (x, y), containing M rows and N columns, where f(x, y), are discrete coordinates. Generally integer values are used for notational clarity for discrete coordinates.

$x = 0, 1, 2\ldots\ldots\ldots(M\text{-}1)$ (1)

$y = 0, 1, 2, \ldots\ldots\ldots(N\text{-}1)$ (2)

So in general, the value of image at any coordinates (x, y) is denoted by f(x, y), where x and y are integers. The section of real plane spanned by coordinates of an image is called the spatial domain, with x and y being referred to as spatial variables or spatial coordinates. We write the representation of an MxN numerical array as

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix} \quad (3)$$

Each element of this matrix is called an image element, picture element, pixel element or pel. The digitization process requires that decisions be made regarding the values for M, N and for number, L, of discrete intensity

levels. *M* and *N* are positive integers. However, due to storage and quantizing hardware considerations, the number of intensity levels. Typically is an integer power of 2.

$L = 2^k$ (4)

The discrete level is equally spaced and that they are integer in the interval [0, *L*-1].The range of values spanned by the gray scale is referred to as dynamic range. Dynamic range of an imaginary system to be ratio of maximum measurable intensity to minimum detectable intensity level in the system. Upper limit of dynamic range is determined by saturation and lower limit by noise.

Generally the number, *b*, of bits required to store a digitized image is:

$b = M*N*K$ (5)

When $M = N$, this equation becomes:

$b = N^2K$ (6)

Eight bits of precision for luminance is common in imaging applications. The eight bit precision is motivated by both the existing computer memory structure 1byte = 8 bits as well as dynamic range of human eye.When an image can have 2^k intensity levels, that image is referred to as *k*-bit image.

TABLE 1 shows the multimedia data types and its requirements.

Table 1: Multimedia Data

Multimedia data	Size/duration	Bits/pixel or Bits/sample	Uncompressed size
Page of text	11” x 8.5”	Varying resolution	16-32 bits
Gray scale image	512x512	8bpp	2.1Mb/img
Color image	512x512	24bpp	6.2Mb/img
Medical image	2048x2048	12 bpp	100 Mb/img
Full motion video	640x640,10	24 bpp	2.21 Gbits

IV. STANDARDS FOR WIRELESS NETWORKS

Wireless networks are among the fastest growing areas in networking research. Wireless communication extends the capabilities of fixed networks to include location independent information storage, transport, retrieval and processing. Future wireless networks allow people on the move to communicate with anyone, anywhere, and at any time using a range of multimedia services and heterogeneous platforms, networks, and devices. Wireless networks play a major role in the application of active networking technology.

Wireless telecommunication history can be classified into different generations of networks. Each generation has been a significant stride which revolutionized the field of mobile or multimedia communication. Era of telecommunication started with 1G in 1980 where all the systems based on analog radio signal technology. Voice was considered to be main traffic. Various 1G standards defined were AMPS, NMT, TDMA and FDMA.In 1990 1G was replaced by 2G which provided rich set of services such as high voice quality and global mobility based on digital radio signal technology. In 2G also voice was main traffic. It includes GSM and GPRS. Both 1G and 2G are based on circuit switched technology for data communication at low speed.2G was huge success. Next 2.5 G is replaced by 3G which includes standards from 2.5 G and also some other technology such as Wi Max. It is based on both circuit switched and packet switched technology providing high data rate with low power consumption. It uses infrastructure of GSM and CDMA to provide its services.

Nowadays, more and more multimedia applications integrate wireless transmission functionalities. Wireless networks are suitable for those types of applications, due to their ease of deployment and because they yield tremendous advantages in terms of mobility of user equipment. However, wireless networks are subject to a high level of transmission errors because they rely on radio waves whose characteristics are highly dependent of the transmission environment. In wireless image transmission applications effective data protection is a crucial issue. JPEG 2000, the newest image representation standard, addresses the issues by including predefined error resilient tools in its encoding system. The main characteristics of JPEG 2000 are: lossy or lossless compression modes; resolution, quality and spatial scalability; transmission and progressive image reconstruction; error resilience for low bit rate applications; region of interest functionality, etc.

4.1 Characteristics of a Wireless Channel

Due to severe wireless channel conditions, such as path loss, fading, co-channel interference, and noise disturbances, the capacity of wireless channels is much lower than wired channels, and bit error rate (BER) is much higher. A wireless network offers both advantages and disadvantages compared to wired network. Advantages of wireless network include mobility and elimination of unsightly cables. Disadvantages of wireless include the potential for radio interference due to weather, other wireless devices, or obstructions like walls. The throughout may fluctuate due to the time varying characteristics of wireless channels. The severe channel conditions have placed another major obstruction when designing efficient image communication systems over wireless networks.

V. IMAGE TRANSMISSION OVER WIRELESS NETWORKS

To design an efficient image communication system over wireless networks[12] there still exist many challenges, of which some are caused by resource limitations, such as power supply, processing capability, some wireless channel conditions, some due to the special characteristics of compressed image data. Image compression is minimizing the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level. The reduction in file size allows more images to be stored in a given amount of disk or memory space. It also reduces the time required for images to be sent over the Internet or downloaded from web pages. There are several different ways in which images can be compressed. For internet use, the two most common compressed graphic image formats are the JPEG format and the GIF format. The JPEG method is more often used for photographs, while the GIF method is commonly used for line art and other images in which geometric shapes are relatively simple. In image compression, a small loss in quality is usually not noticeable. When there is some tolerance for loss, the compression factor can be greater than it can when there is no tolerance. For this reason, graphic images can be compressed more than text files or programs. Since image data contains a lot of redundancy, to efficiently utilize limited resources, source compression is always necessary. Wavelets offer an elegant technique for representing the levels of details present in an image. When an image is decomposed using wavelets, the high pass component carry less information. And low pass components carry more information. The possibility of elimination of high pass components gives higher compression ratio in wavelet based image compression. To achieve higher compression ratios, various coding schemes have been used. Some of the well-known coding algorithms are EZW (Embedded Zero tree Wavelet),SPIHT(Set Partitioning inHierarchical Tree) and EBCOT(Embedded Block Coding with optimal Truncation).SPIHT has been one of the popular schemes used for compression. The SPIHT algorithm, developed by Said and Pearlman

in 1996 is a fast and efficient image compression algorithm works by testing ordered wavelet coefficients for significance in a decreasing bit plane order, and quantizing only the significant coefficients.

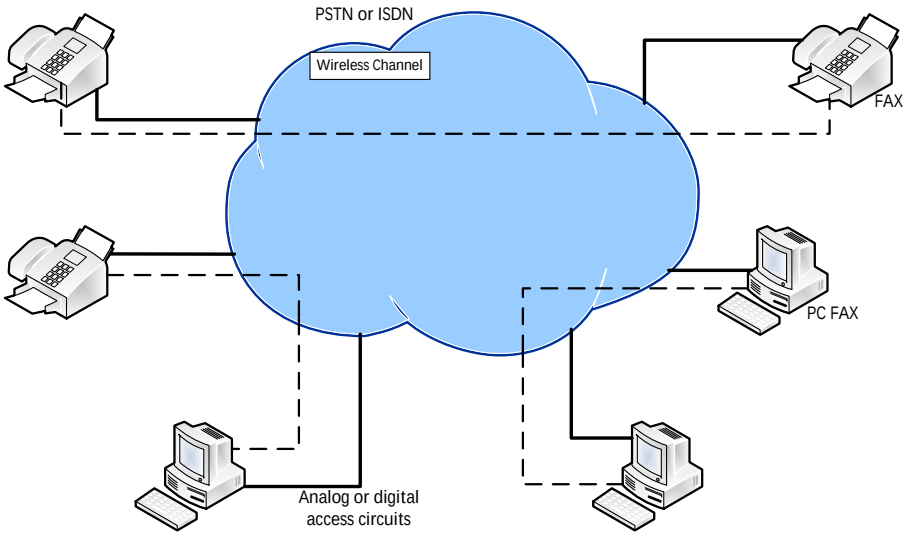


Figure 1: A Wireless Network Schematic

The high coding efficiency obtained by this algorithm is due to group testing of the coefficients of a wavelet tree. The SPIHT algorithm is a refined version of EZW algorithm. It can perform better at higher compression ratios for a wide variety of images than EZW. The algorithm uses a portioning of trees in a manner that tends to keep insignificant coefficients together in larger subsets.

The SPIHT algorithm groups the wavelets coefficients and trees into sets based on their significant information. The encoding algorithm consists of two main stages, sorting and refinement. In the sorting stage, the threshold for significance is set to as 2n, where n is the bit level, and its initial value is determined by number of bits required to represent the wavelet coefficient with maximum absolute value. Significance for trees is obtained by checking all the member detail coefficients.

In Compared with general data, the compressed image data has some special characteristics, such as unequal importance, error tolerance and constrained error propagation. In this, we consider wireless image transmission system for images coded using SPIHT with baseline JPEG 2000and 3G as the medium and propose a FEC scheme based on low density parity check codes.

When an image is transmitted from source to destination over wireless links it will be subject to random bit errors caused by channel impairments. A single bit error will result on the occurrence of a corrupted of neighboring bits also. This stops the image transmission process and whole process has to be repeated again and again until there is no loss. This results in incomplete transmission. An efficient image transmission system over wireless networks is faced by many challenges, of which some are caused by resource limitations such as:

- Wireless channel conditions
- Special characteristics of compressed image data. The characteristics of compressed data is its PSNR value and compression ratio.The PSNR (dB) is calculated with the following formula:

$$PSNR = 10\log_{10}\left[\frac{MaxgreyLevel*MN}{\sum_{xy}|g(x,y)-f(x,y)|}\right] \quad (7)$$

Where g(*x*,*y*) is the compressed image,f(*x*,*y*) is the raw image,*M* is the image width, *N* is the image height and max grey level is the maximum value of f(*x*,*y*).A maximum grey level is equal to 255.

· Maintaining a good quality of service (QoS) for multimedia over wireless networks

Multimedia over wireless will certainly need a higher bandwidth. As wireless data transmissions incur the most data loss and distortion, error resilience and error correction become primary concerns. Wireless channels have more interference than wired channels, with specific loss patterns depending on the environment conditions. The bitrates for wireless channels are also much limited as 3G bit rates are most suitable for images. This implies that although a lot of bit protection must be applied, coding efficiency has to be maintained as well. Error resilient coding is important. The QoS parameters specified for wireless multimedia data transmission depends on the following parameters:

· Data rate: A measure of transmission speed, expressed inkilobits per second (kbps) or megabits per second. (Mbps)

· Latency (maximum frame/packet delay): Maximum time needed from transmission to reception, measured in milliseconds.

· Packet loss or error: A measure of error rate of the packetized data transmission. Packets get lost or garbled, such as over the internet. They may also be delivered late or in wrong order. Since retransmission is often undesirable, a simple error recovery method for multimedia is to repay the last packet, hoping the error is not noticeable.

· In general, for uncompressed image desirable packet loss is $<10^{-2}$.For compressed multimedia data desirable packet loss is less than 10^{-7} to 10^{-8}.

· Jitter (or delay jitter): Jitter is referred to variance of frame/packet delays. It is the worst case of variation in delay.

· Sync skew: A measure of multimedia data synchronization, often measured in milli-seconds. In general, ±200 msec is acceptable.

All types of multimedia information are stored and processed within a computer in adigital form.

Table 2: Comparison Of Various Networks In Terms Of User Requirements

Network Generation	Sound	Sight	Knowledge
1G-2G	Voice	--	Low speed Data
3G	Voice	Images	Hyper text
4G	Voice, speech	Video	Files (speech, HT, video)
Typical Band width	10-80 kbps	1-20Mbps	0.5-10Mbps
Required Latency	<160ms	<100ms	<5s
Principle Application	Communication	Entertainment	Information

VI. TURBO CODES – THE CURRENT CHANNEL CODING ALGORITHM FOR 3G

7.1 ABOUT 3G

Wireless services have the highest demand in wireless and internet world. Every generation technology has some plat form for its development.1G was based on analog signaling whereas 2G on low-band digital data signaling. The 3G technology was developed to overcome the faults of 1G and 2G.3G finds its applications in wireless technologies, as high speed transmission, advanced multimedia access .It can support data rates between 128 and 144 kbps for devices that are moving fast and 384 kbps for slow ones. For fixed wireless LANs, the speed goes beyond 2mbps.The spectral efficiency ranges from 1-5Mhz.3G supports packet switching and its network architecture is wide-area cell based [19].

3G technology provides both circuit design and Packet design. Circuit design, being the oldest, has greater ability to hold connection for a longer duration. On the other hand the packet design is a wireless technology and is the core part of internet data transmission. The combination of these two patterns helps 3G technologies to perform better and faster.3G uses licensed spectrumin 3G, network based QoS depends on following to provide a satisfactorily service as: Throughput, Packet Loss Rate, reliability and delay.3G can be used with wireless LAN for better quality of service.

Channel coding is the key element of any digital wireless communication systems [21] since it minimizes the effect of noise and interference of any transmitted signal. In 3G wireless systems channel coding techniques must serve both voice and data users whose requirements considered may vary. Thus in third generation partition project (3GPP) standard offers two coding techniques convolutional coding for voice and Turbo-coding for data services.

Channel coding allows to reduce the power by maintaining the QoS or to improve QoS for a given transmitted power.3G systems offers different data rates and coding techniques to satisfy the varying latency, throughput and error-performance requirements. Example data services which require lower error rate and high throughput, but can tolerate a larger latency.

7.2 About Turbo Codes

The implementation complexity of encoders in turbo coding is negligible. Both are trellis encoders which map a long input sequence to a coded data stream. They consist of shift registers and an interleaved address generator. The decoders are based on trellis propagation of received input sample sequence to calculate a maximum likelihood sequence or bit detection.

The basic system consists of two identical systematic recursive convolutional encoders connected in parallel with an inter-leaver preceding the second recursive convolutional (RSC) encoder [20]. RSC encoders encode the information bits. The first encoder operates on the input bits in their original order while the second one operates on the input bits as permuted by the inter-leaver.

Forward error correction is enabled by introducing parity bits. For turbo codes, the original information is denoted as systematic information is transmitted together with the parity information. The source for 3GPP consist of two recursive systematic encoders with constraint length k=4.Each transmitted block is iteratively decoded. The systematic information and [parity information serves as input to first component to decoder. The decoding algorithm involves the joint estimation of two Markov processes one for each constituent code.

The 3G standard specifies the encoder structure and parameters like block size and throughput requirements. The maximum block size of turbo codes including the data, frame quality indicates CRC and two reserved bits is set to 5114.During encoding an encoder output tail sequence is added which appends another 3 bits the tail bits, for both systematic and parity information of each encoder.

VIII. LDPC CHANNEL CODING FOR 3G

We propose to use LDPC coding as an alternative algorithm for 3G channel coding. LDPC codes were invented by Gallager in early sixties, [25] their importance as capacity approach and performance analysis issues are currently being addressed in coding environments and applications. Basically LDPC codes are linear block codes with a very sparse parity check matrix H order NxM. Typically, the matrix H is generated by applying random perturbations to the zero matrixes until a specified number of ones appear in each column and roughly

fixed equal number of ones appear in each row. The associated generator matrix *G* can be obtained by Guassian elimination of *H*, where *G* is not necessarily sparse. The pseudo random parity check matrix also leads to LDPC codes that have random like properties with channel coding theorem conditions.LDPC codes could be used to transmit information reliably at the rates close to channel capacity. The decoding can be done by sum product algorithm. This is an iterative probabilistic decoding algorithm that begins with *H* and a set of prior probabilistic for the *N* bits. It then iteratively updates these on *M* parity checks until all of the parity checks are satisfied. Sometimes decoder fails if it cannot satisfy all parity checks. Decoding algorithm works well provided *H* does not contain patterns where two columns have two or more check positions in common.

The advantages of using this coding algorithm are as follows [9]:

- LDPC codes with iterative decoding provide very good performance over a variety of channels with reasonably low complexity. In particular irregular LDPC codes outperform turbo codes and regular codes and to approach to channel capacity of several channels at large block lengths.
- Another advantage of LDPC codes over turbo codes is efficient hardware implementation of the decoder.
- These algorithms are parallelizable and can be realized at much faster speed than turbo decoders and thirdly almost all errors are detectable []. This is because in turbo coding a large decoding delay is introduced owing to large block lengths and many iterations of decoding required for near-capacity performance and significantly weakened performance at BERs below 10^{-5}.

Due to above advantages of LDPC codes we have chosen these codes for encoding and decoding purpose for wireless image transmission maintaining quality of service .

IX. CONCLUSION

In this paper, we present an overview of the image transmission process over wireless networks. We describe the various techniques for source coding, channel coding and transmission of images through wireless networks. We find that there are research opportunities that exist in channel coding since there are significant advantages of using LDPC codes over Turbo codes.

REFERENCES

[1] ApurvaSinha ,MukeshKumar,A.K. Jaiswal, RohiniSaxena,"Performance , Analysis of High Resolution Images Using Interpolation Techniques in Multimedia communication System", Signal and Image processing Journal, Vol.5, No.2, April 2014.

[2] Feng Cen,"Distributed Joint Source andChannel Coding with Low-Density parity Check codes", IEEE Communications letters, VOL.17, No.12, December 2013.

[3] Zheng Guo, JieHuang,BingWang,"A practical Joint Network-Channel coding Scheme for reliable communication in Wireless networks", IEEE transactions on Wireless Communications, VOL.11, No. 6, June 2012.

[4] Yang Hu; Pearlman, W.A; Xin Li,"Progressive Significance MAP and its applications", IEEEE transactions on Image Processing Vol. 21, Issue 7, 2012.

[5] Chien-Chen Lin,YaoLiand Chen-Yi Lee," A predefined Bit-plane Comparison Coding for Mobile video applications ", IEEE transactions on Circuits and Systems Vol.58,No.7 July 2011.

[6] Zrae RA, Hassan M, EI Tarhuni M, "Wireless image transmission using joint adaptation of modulation and channel coding", Computers and Communication (ISCC), IEEE Symposium, 2011.

[7] Maria Fresia, Fernando Perez-Cruz, H. Vincent Poor, SergisVerdu, "Joint Source and Channel coding proposing a low density parity-check code", IEEE signal processing Magazine November 2010.

[8] Max AGUEH,"Wireless Multimedia Communications and Networking based on JPEG 2000", Communications and Networking book.ISBN 978-953-307-114-5, pp.434, September 2010, Sciyo,Croatia.

[9] Abraham Gabay, Michel Kieffer, Pierre Duhamel, "Joint Source-Channel coding using Real BCH Codes for Robust Image transmission", IEEE transactions on Image processing,Vol.16,No.6,June 2007.

[10] Pan, Xiang, Amir H. Banihashemi, and Cuhadar, Aysegul, "Progressive Transmission of Images over Fading Channels Using Rate-compatible LDPC codes", IEEE Transactions on Image Processing, Vol.15, No. 12, December 2006.

[11] Xiang Pan,AysegulCuhadar,Member, IEEE and Amir H.Banihashemi Senior Member, IEEE "Combined Source and Channel coding with JPEG 2000 and rate compatible Low-Density Parity Check codes",IEEE Transactions on Image Processing, Vol.15, No. 12, December 2006.

[12] Nikolaos Thomas, NikoloasV.Boulgouris, Michael G.Strintzis,"Optimized Transmission of JPEG2000 streams over wireless Channels",IEEE Transactions on Image Processing,Vol.15.No.1.January 2006

[13] Abdullah Al Muhit and Teong Chee Chuah," Robust Quality – Scalable Transmission of JPEG2000 Images over Wireless Channels using LDPC codes", ISCV pp.28-39, Springer-VerlagBerlin,Heidelberg 2006

[14] Y Sriraja, Tanja Karp, Sunanda Mitra, "Error-resilient wavelet coding for wireless image transmission", Dept. of Electrical and Computer Engg., Texas Tech University, 2005.

[15] Aria Nosratinia, Jin Lu, BehnaamAazhang,"Source – Channel Rate Allocation for Progressive Transmission of Images", IEEE Transactions on Communications, Vol. 51, No.2, February 2003.

[16] B.A.Banister, B.Belzer and T.R Fischer,"Robust Image Transmission using JPEG 2000 and turbo codes", IEEE Signal process. Lett. Vol.9, no.4, pp.117-119, Apr.2002.

[17] Amir Said, and William A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchal Trees", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 6, No.3, June 1996.

[18] Shapiro, Jerome M, "Embedded Image Coding using Zero-trees of Wavelet Coefficients", IEEE Transactions in Signal Processing, Vol. 41, No. 12, December 1993.

[19] D. Taubman and M. Mercellin, JPEG2000: Image Compression Fundamentals, Standards and Practice, Norwell, MA: Kluwer, 2001.

[20] Shu Lin and Daniel J. Costello, Jr, Error Control coding, Second Edition by Pearson Education, Inc, Publishing as Prentice Hall, Copyright 2005.

[21] Dr. Sunil Kumar S. Manvi, Mahabaleshwar S. Kakkasageri, Wireless and Mobile Networks concepts and Protocols, Wiley India Pvt. Ltd, 2010.

[22] Rafael C. Gonzalez, Richards E. Woods, Digital Image Processing, Third Edition, Pearson Prentice Hall, Copyright 2008.

[23] FriedbertBerens,GerdKreiselmaier, Norbert When, "Channel Decoder Architecture for

[24] 3G Mobile Wireless Terminals",Design, Automation and Test in Europe Conference and Exhibition, 2004. Proceedings, 16-20 Feb. 2004, 192 - 197 Vol.3.

[25] R.G.Gallagher,"Low Density Parity Check Codes", MIT Press, Cambridge, MA, 1963.

DRIVER BEHAVIOR ANALYSIS USING NON-INVASIVE SENSORS

[1]L. Nikitha , [2]J.Kiranmai, [3]B.Vidhyalakshmi

[1,2,3]Department of Electronics and Communication, SCSVMV University, Kanchipuram , (India)

ABSTRACT

Detecting driver distraction has been an important research topic over the past few years. While there are many common reasons for vehicle crashes, driver distraction and inattention are very prominent causes. Previous studies have shown the impact that inattentions can have on driving behavior, which can lead to many crashes and fatalities. The study reported by The National Highway Traffic Safety Administration (NHTSA) indicated that over 25% of police-reported crashes involved inattentive drivers. The 100-car Naturalistic Study concluded that over 65% of near crashes and 78% of crashes included inattention. These high percentages are not surprising, since 30% of the time drivers are in a moving vehicle, they are engaged in secondary tasks that are potentially distracting. These numbers are estimated to increase as the usage of in-vehicle technologies for navigation, communication and infotainment, and the number of cars on the roads is expected to exponentially increase in the next years. A driver monitoring system that is able to sense inattentive drivers can play an important role in reducing the number of accidents, preventing fatalities and increasing the safety on roads. In this paper Monitoring driving habits of a person is presented, which is based on CAN network. The CAN bus is used as a communication of a distributed control network. This paper mainly introduces the design of the hardware and the software in detail. This device track speed, Engine temperature and Alcohol consumption status.

I INTRODUCTION

There has always been a rapid development of technology in every field. In past few years with the implementation and fast development of the in-vehicle technologies especially in the cars had lead to the vast increase in the unintentional accidents. These accidents are caused by the distraction in the driver's attention due to the involvement in the secondary tasks like operating the radio, GPS, and other in-vehicle technologies. The driver distractions are detected using the various non-invasive sensors. The driver is made to involve in the secondary tasks like operating on the radio, navigation systems like GPS(Global Positioning System) and mobile phones, etc. The changes in the driver's behavior is detected by comparing the values obtained by the sensors when the driver is involved in secondary tasks and when he is not involved in the secondary tasks.

When any abnormality is detected in any of the sensors then the safety measures like indication by the side lights, an alarm system, gradually closing the valve of the fuel tank and the opening of the airbag are taken.

The paper is organized as follows: Section 2 presents the open challenges currently existing in the real time and the related works done in this field of analyzing the driver's behavior. Section 3 describes the methodology of

the data collection and the working procedure of the system. Section 4 studies the effects in driver behavior induced by the secondary tasks, including the statistical analysis of multimodal features and discriminative analysis between normal and task driven conditions. Section 5 concludes the paper with final remarks, limitations of the study and the further research directions.

II RELATED STUDIES

Secondary tasks deviate the driver's attention from the primary driving task [1]. Various activities have been proposed to induce cognitive and/or visual distractions. For cognitive distractions, common approaches include solving math problems [2], [3], [4], [5], talking to another passenger [1], [6], and focusing on other activities such as following the stock market [5]. Common secondary tasks for visual distraction are "look and find" tasks [7], [3], [8], operating devices such as a touchscreen [1], or a cellphone [9], and reading sequences of numbers [2]. While these cognitive and visual tasks clearly affect the driver, some of them may not represent the common distractions observed in real scenarios.

While most of the studies on driver behaviors rely on simulators [10], [11], [3], [1], [8], some studies have considered recordings in cars equipped with multiple sensors [2], [12], [13], [14], [6], [15]. Perez et al. [12] presented the "Argos" system for data collection. Murphy-Chutorian and Trivedi [13] reported results on data recorded in the LISA-P experimental testbed. The car has video and motion cameras with near- IR illuminator[. They have used computer visual algorithms to automatically extract visual information, achieving promising results towards detecting driver distraction. Another data collection vehicle was designed by Takeda et al. [16]. The car is equipped with cameras and microphones, laser scanners (front, back), pedal pressure and physiological sensors. A similar car was designed by Abut et al. [17] called UYANIK. The UTDrive is another car platform,[18], [19]. These cars provide more realistic data to study driver behaviors.

Frontal cameras can be useful to assess the distraction level of the driver [13], [20]. Relevant visual features include head pose, gaze range and eyelid movements [2], [7], [8],[21], [22]. Liang et al. [7] showed that eye movements and driving performance measures were useful for detecting cognitive distraction. Su et al. [21] presented an approach to monitor visual distractions using a low cost camera. The study relied on eyelid movements and face orientation to predict driver's fatigue and distraction. Azman et al. [22] used eye and lip movements to predict cognitive distractions in simulated environment. Kutila et al. [2], [4] extracted gaze angle, head rotation and lane position for cognitive distraction detection. Bergasa et al. [14] proposed to predict fatigue with percent eye closure (PERCLOS), eye closure duration, blink frequency, face position, fixed gaze and nodding frequency. They used IR illuminator to mitigate the changes in illumination. A similar approach was presented by Zhu and Ji[23]. Other studies have considered cameras for capturing and modeling foot gestures for brake assistance systems [24], [25],[26].

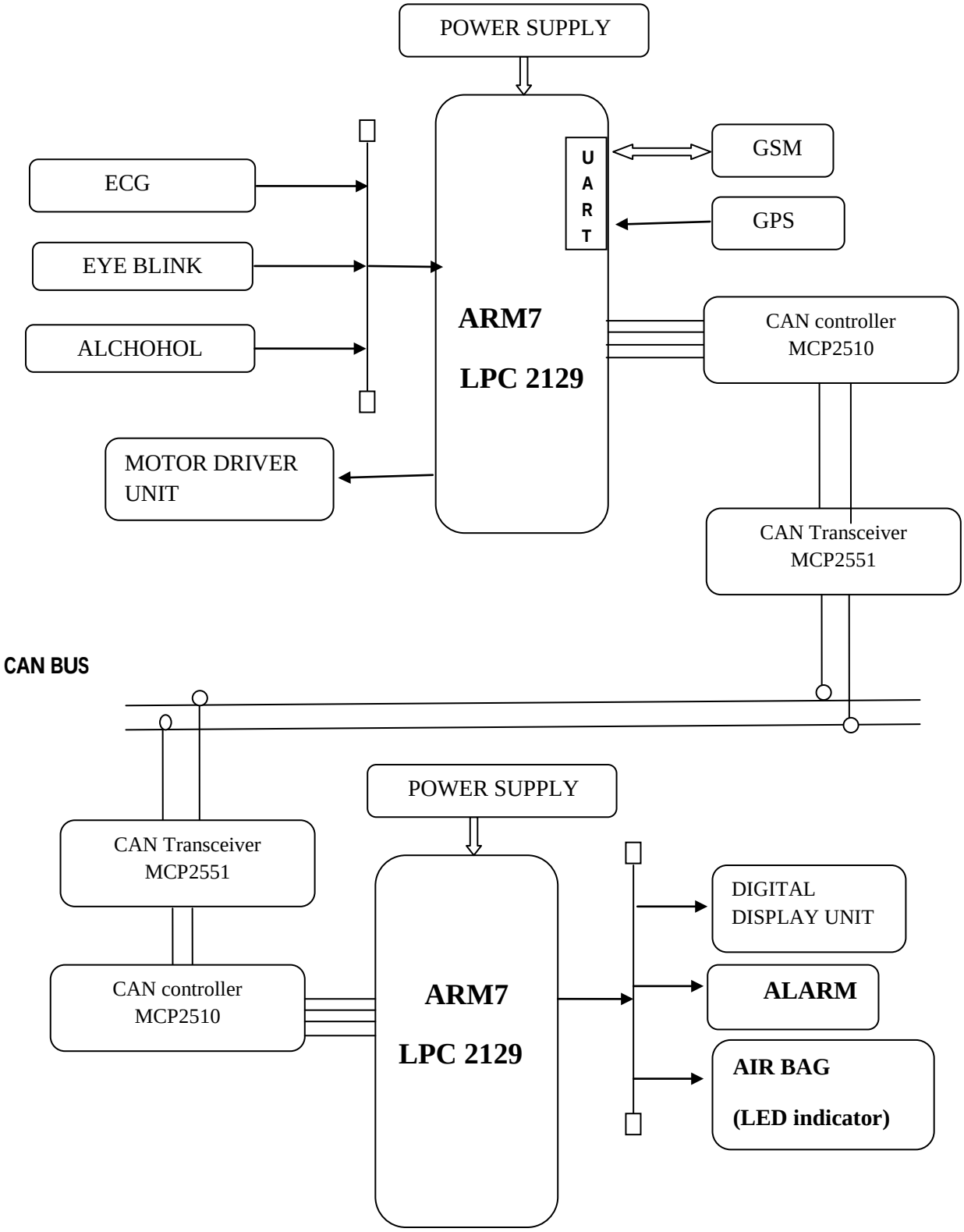


Fig 1: Working Modality of the Module

III METHODOLOGY AND WORKING MODALITY

A real-world driving study inherently involves numerous uncontrollable variables such as traffic, weather and traffic lights which are not easily replicated in a simulated environment. This analysis aims to identify relevant features that are directly affected by the driver's behaviors, and to use these features to quantify the distraction level of the drivers.

Before the motor can start the driver has to give the alcohol test. If any alcohol content is detected then the valve of the fuel tank will be closed. When no alcohol content is detected then the motor starts. While driving the sensors like eye blink and pulse sensors will monitor the eye blink and heart beat continously. If any abnormalaties are detected then the accident preventive measures like the opening of the air bag, sounding the alarm system and indication using the side lights are taken. The vibration due to the crash by the other vehicle can be detected using the vibration sensor and if any vibration due to the crashes is detected, then the accident preventive measures are taken.

With the help of the CAN(Controller Area Network Bus the signals from a controller which monitors the driver's behavior and to which the sensors are interfaced is communicated to the other controller to which the external features are connected. In case the accident occurs then the message about the accident is given through the GSM technique. The position of the car is tracked using GPS (GLOBAL POSITIONING SYSTEM).

The output of the heart beat sensor is active high for each and every heartbeat and this can be given to the microprocessor. LED flashes for each heartbeat whenever sensor starts working. The sensor works on the basis of light modulation by blood flow through finger for each pulse.

Fig 2: ECG Sensor

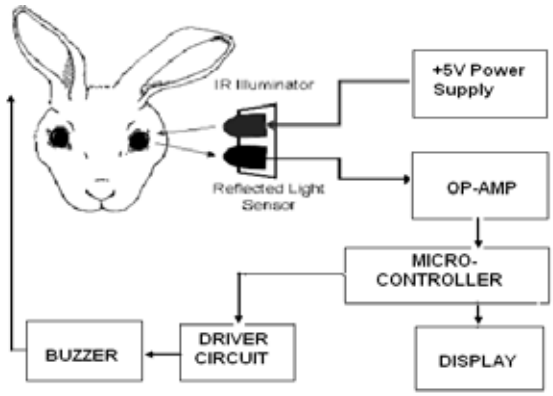


Fig 3: Eye blink sensor

The EBM (Eye Blink Monitoring) system works on the principle of monitoring eyelid movements of the person inside the vehicle continuously by the use of an IR sensor. Whenever the driver falls asleep, then the buzzer will rang to wake him up.

The Alcohol breathalyzer estimates the Blood Alcohol Concentration by detecting and measuring the presence of the ethanol vapour in our breath. The alumina tube and the coils are the heating system, the yellow, brown

parts and the coils in the picture. If the coil is heated up SnO2 will become the semi conductor further we can find more movable electrons, which indicates the current flow is ready. Whenever the alcohol gas molecules meet the electrode which is present in the air, then gradually ethanol in the air burns into acetic acid which further produces more current.

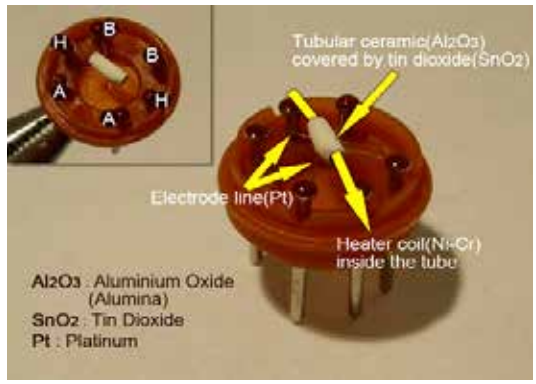


Fig 4: Gas sensor

IV EFFECTS IN DRIVER BEHAVIOR

The secondary tasks force the driver to be distracted from the primary task which is driving. These distractions in the driver behavior can be detected by using a series of sensors.

The normal pulse rate of the human being is about 72 heart beats per second. In case of any stroke or any other problems related to the heart beat then the count of the heart beat varies from the normal value. The range from 65 to 80 heart beats is considered to normal. If the pulse rate is not in this range then the driver's behavoir changes due to the abnormalities.

The normal eye blink rate is about 15 to 20 blinks or lesser. When the driver is in drowsy state the number of blinks may increase than the normal level as the driver tries to come out of the drowsy state or the eye lids remain closed for a longer period than it is required for the normal eye blink.

The breath of the person who is not drunk does not contain any alcohol content or very minute traces of the alcohol. If the breath contains more than the allowed level of the alcohol content then it is concluded that the person is in drunken state.

V CONCLUSION

The system proposed in the paper is a fully functional automatic mentoring system which is a step ahead of presently available in terms efficiency, accuracy and simplicity. The system is designed with an objective of minimizing road accidents that occur due to over speeding. The algorithm used for position matching and subsequent speed limit extraction is the first of its kind relying on both GSM and GPS input signals, complementing each other and thus avoiding limitations of using them individually for position tracking.

This system can be implemented in the larger vehicles like trucks and buses and further more vehicles like passenger and cargo trains. This proposed system can be used as the warning system to avoid collisions in the National Highways.

This system can be further developed using the image processing technique to analyze the driver's facial expressions to understand the different situations that can arise while driving.

REFERENCES

Proceeding papers

[1] T. Ersal, H. Fuller, O. Tsimhoni, J. Stein, and H. Fathy, "Model-based analysis and classification of driver distraction under secondary tasks," IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 3, pp. 692-701, September 2010.

[2] M. Kutila, M. Jokela, G. Markkula, and M. Rue, "Driver distraction detection with a camera vision system," in IEEE International Conference on Image Processing (ICIP 2007), vol. 6, San Antonio, Texas, USA, September 2007, pp. 201-204.

[3] F. Putze, J.-P. Jarvis, and T. Schultz, "Multimodal recognition of cognitive workload for multitasking in the car," in International Conference on Pattern Recognition (ICPR 2010), Istanbul, Turkey, August 2010.

[4] M. Kutila, M. Jokela, T. Makinen, J. Viitanen, G. Markkula, and T. Victor, "Driver cognitive distraction detection: Feature estimation and implementation," Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 221, no. 9, pp. 1027-1040, September 2007.

[5] J. Harbluk, Y. Noy, P. Trbovich, and M. Eizenman, "An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance,"Accident Analysis and Prevention, vol. 39, no. 2, pp. 372-379, March 2007.

[6] A.Sathyanarayana, P. Boyraz, Z. Purohit, R. Lubag, and J. Hansen, "Driver adaptive and context aware active safety systems using CANbus signals," in IEEE Intelligent Vehicles Symposium (IV 2010), San Diego, CA, USA, June 2010.

[7] Y. Liang, M. Reyes, and J. Lee, "Real-time detection of driver cognitive distraction using support vector machines," IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 2, pp. 340-350, June 2007.

[8] F. Tango and M. Botta, "Evaluation of distraction in a driver-vehicleenvironment framework: An application of different data-mining techniques," in Advances in Data Mining. Applications and Theoretical Aspects, ser. Lecture Notes in Computer Science, P. Perner, Ed. Berlin, Germany: Springer Berlin / Heidelberg, 2009, vol. 5633, pp. 176-190.

[9] J. Jain and C. Busso, "Analysis of driver behaviors during common tasks using frontal video camera and CAN-Bus information," in IEEE International Conference on Multimedia and Expo (ICME 2011), Barcelona,

Spain, July 2011.

[10] "EEG-based drowsiness estimation for safety driving using independent component analysis," IEEE Transactions on Circuits and Systems I:Regular Papers, vol. 52, no. 12, pp. 2726-2738, December 2005.

[11] I. Damousis and D. Tzovaras, "Fuzzy fusion of eyelid activity indicators for hypovigilance-related accident prediction," IEEE Transactions on Intelligent Transportation Systems, vol. 9, no. 3, pp. 491-500, September 2008.

[12] J. Zamorano, "Argos: An advanced in-vehicle data recorder on a massively sensorized vehicle for car driver behavior experimentation," IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 2, pp. 463–473, June 2010.

[13] E. Murphy-Chutorian and M. Trivedi, "Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness," IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 2, pp. 300-311, June 2010.

[14] L. Bergasa, J. Nuevo, M. Sotelo, R. Barea, and M. Lopez, "Realtime system for monitoring driver vigilance," IEEE Transactions on iIntelligent Transportation Systems, vol. 7, no. 1, pp. 63-77, March 2006.

[15] A. Sathyanarayana, S. Nageswaren, H. Ghasemzadeh, R. Jafari, and J.H.L.Hansen, "Body sensor networks for driver distraction identification," in IEEE International Conference on Vehicular Electronics and Safety (ICVES 2008), Columbus, OH, USA, September 2008.

[16] K. Takeda, J. Hansen, P. Boyraz, L. Malta, C. Miyajima, and H. Abut, "International large-scale vehicle corpora for research on driver behavior on the road," IEEE Transactions on Intelligent Transportation Systems, vol. 12, no. 4, pp. 1609-1623, December 2011.

[17] H. Abut, H. Erdo˘gan, A. Ercil, B. Curuklu, H. Koman, F. Tas A. Arguns¸ah, S. Cos¸ar, B. Akan, H. Karabalkan et al., "Real-world data collection with "UYANIK"," in In-Vehicle Corpus and Signal Processing for Driver Behavior, K. Takeda, H. Erdo˘gan, J. Hansen, and H. Abut, Eds. New York, NY, USA: Springer, 2009, pp. 23-43.

[18] P. Angkititrakul, M. Petracca, A. Sathyanarayana, and J. Hansen, "UTDrive: Driver behavior and speech interactive systems for in-vehicle environments," in IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, June 2007, pp. 566-569.

[19] P. Angkititrakul, D. Kwak, S. Choi, J. Kim, A. Phucphan, A. Sathyanarayana, and J. Hansen, "Getting start with UTDrive: Driver-behavior modeling and assessment of distraction for in-vehicle speech systems," in Interspeech 2007, Antwerp, Belgium, August 2007, pp. 1334-1337.

[20] M. Trivedi, T. Gandhi, and J. McCall, "Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety," IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 1, pp. 108-120, March 2007.

[21] M. Su, C. Hsiung, and D. Huang, "A simple approach to implementing a system for monitoring driver inattention," in IEEE International Conference on Systems, Man and Cybernetics (SMC 2006), vol. 1, Taipei, Taiwan, October 2006, pp. 429-433.

[22] A. Azman, Q. Meng, and E. Edirisinghe, "Non intrusive physiological measurement for driver cognitive distraction detection: Eye and mouth movements," in International Conference on Advanced Computer Theory and Engineering (ICACTE 2010), vol. 3, Chengdu, China, August 2010.

[23] Z. Zhu and Q. Ji, "Real time and non-intrusive driver fatigue monitoring," in IEEE International Conference on Intelligent Transportation Systems, Washington, DC, October 2004, pp. 657-662. IEEE TRANSACTIONS ON MULTIMEDIA, VOL. X, NO. X, APRIL 2012 13

[24] J. McCall and M. Trivedi, "Driver behavior and situation aware brake assistance for intelligent vehicles," Proceedings of the IEEE, vol. 95, no. 2, pp. 374-387, February 2007.

[25] C. Tran, A. Doshi, and M. Trivedi, "Modeling and prediction of driver behavior by foot gesture analysis," Computer Vision and Image Understanding, vol. 116, no. 3, pp. 435-445, March 2012.

[26] C. Berka, D. Levendowski, M. Lumicao, A. Yau, G. Davis, V. Zivkovic, R. Olmstead, P. Tremoulet, and P. Craven, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," Aviation, Space, and Environmental Medicine, vol. 78, no. 5, pp. 231-244, May 2007.

[27] Detection of Intoxicated Drivers Using Online System Identification of Steering Behavior Mehran M. Shirazi, and Ahmad B. Rad, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 15, NO. 4, AUGUST 2014.

Journal papers:

[28] Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel Sang-Joong Jung, Heung-Sub Shin, Wan-Young Chung, Published in IET Intelligent Transport Systems , Received on 20th February 2012 , Revised on 16th July 2012 ,Accepted on 18th July 2012, doi: 10.1049/iet-its.2012.0032.

INTELLIGENT INTRUSION DETECTION SYSTEM IN WIRELESS SENSOR NETWORKS

S.Yamunarani[1], D.Sathya[2], S.Pradeepa[3]

[1]*Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore, TamilNadu, (India)*

[2]*Assistant Professor, Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore, TamilNadu, (India)*

[3]*Assistant Professor, Department of Information Technology , Adithya Institute of Technology, Coimbatore, TamilNadu, (India)*

ABSTRACT

Wireless Sensor Networks are highly distributed networks of tiny, light-weight wireless nodes, deployed in large numbers to monitor the environment. Monitoring the system includes the measurement of physical parameters such as temperature, pressure, relative humidity and co-operatively passing their data to the main location. Intrusion Detection System can act as a second line of defense and it provides security primitives to prevent attacks against computer networks. The Intelligent Intrusion Detection System has the capability to do different things without any human intervention. Intrusion Detection System uses Misuse-based and Anomaly-based methods. Intelligent Intrusion Detection System technique such as Back Propagation Network, Bayesian Belief Networks and Support Vector Machines are used for detecting cross layer attacks to reduce the false positive rate and improve detection accuracy.

Keywords: Intelligent Intrusion Detection, Wireless Sensor Networks, Back Propagation Networks, Bayesian Belief Networks, Support Vector Machines.

I. INTRODUCTION

The Wireless Sensor Networks [WSNs] consists of sensor nodes ranging from few hundred or even thousands depending on the application. Each sensor node may be connected to one or more other sensor nodes. Each node of the sensor network consists of four components, namely, sensing unit, processing unit, transceiver unit and the power unit which is shown in Fig 1. In addition to the above units, a wireless sensor node may include a number of application-specific components. For this reason, many commercial sensor node products include expansion slots and support serial wired communication.

Detection based techniques aim at identifying the intrusions that affect the network infrastructure after a failure of the prevention based techniques. In our work it focused on Intrusion Detection System [IDS] applied to wireless sensor networks. The two major models of Intrusion Detection include Anomaly detection and Misuse detection. Anomaly detection builds a model of normal behavior and compares the model with detected behavior. From [5], Anomalydetection has a high detection rate, but the false positive rateis also high.

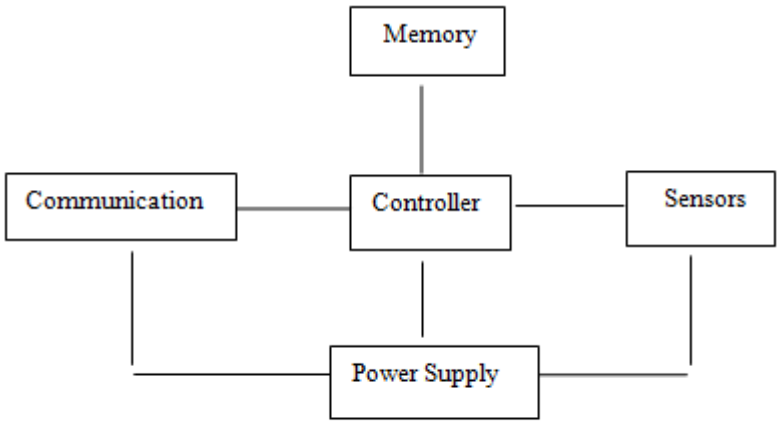


Fig 1. Components of Sensor Node

The advantage of Anomaly detection has the ability to detect unknown attacks by using anomaly IDS. The Disadvantage is, it cannot detect different types of attacks. The Misuse detection model is built, so that theattack type is determined by comparison with the attack behavior. The advantage of Misuse detection has high accuracy, but the detection rate is lower. The Disadvantage of Misuse detection cannot detect unknown attacks in the networks.

The remainder of this paper is organized as follows: InSection 2, works relevant to the common attacks in WSN and the analytic tools of intrusion detection, used in our research, are introduced. In Section 3, the Existing methods are discussed. The simulation results used to evaluate the performance of the existing system are presented in Section 4. The proposed system is discussed. Finally, the conclusion and future work are discussed in Section 5.

II. RELATED WORK

In [4], IDS is the second line of defense and it gives security to prevent attacks against computer network. Hybrid composed of central agent perform accurate IDS by using Data mining techniques. In proposing a hybrid system, lightweight, distributed Intrusion Detection System for wireless sensor networks are used. This IDS uses both Misuse-based and Anomaly-based detection techniques. The compared techniques are Classification And Regression Tree(CART), Chi-squared Automatic Interaction Detection(CHAID), C5.0, Logistic Regression, BayesianNetwork. According to the experimental results, the best detection techniques are C5.0 and CART. The advantages of these two techniques are more accurate, and they also show lower false positive rate, that implies low energy consumption for alert communication from Local Agents to the Central Agent.

In [7], sink and cluster head are easily attacked by enemies, so the security is necessary. The capabilities of all sensors in Cluster-Based WSN are heterogeneous. Due to different capabilities and probabilities of attack on them, three separate IDS are designed to sink, Cluster head and sensor nodes. Intelligent Hybrid IDS[IHIDS] is proposed, it has learning ability. Hybrid IDS is proposed by cluster head, it is same as IHIDS but it has no learning ability.

The Anomaly detection model is used as the first line of defense in IHIDS. The anomaly detection model acts like a filter. BPN learns the corresponding relations between input and output variables, and tunes the corresponding weight. The advantage of BPNis a simple and fast detection method is used in the sensor node to avoid overwork and save resources for the purpose of safety.

In [9], HIDS consists of an Anomaly detection model and Misuse detection model and decision making model. Anomaly model uses Rule based method. It filters a large number of packet records, using the Anomaly detection model, and performs a second detection with the Misuse detection model, when the packet is determined to intrusion. It efficiently detects intrusion, and avoids the resource waste. Misuse detection uses Back Propagation network.

Finally, it integrates the outputs of the Anomaly detection and Misuse detection models with a decision making model. The output of the decision making model is then reported to an administrator for follow-up work. This method not only decreases the threat of attack on the system, but also helps the user handle and correct the system further with hybrid detection. The advantage of decision making model is simple and fast. The proposed model lowers false positive rate and achieves a high detection rate and accuracy.

In [1],WSNs are susceptible to some types of attacks since they are deployed in open and unprotected environments and are constituted of cheap small devices. Preventive mechanisms can be applied to protect WSNs against some types of attacks. There are some attacks for which there is no known prevention methods, such as wormholes.

Besides preventing the intruder from causing damage to the network, the Intrusion Detection System (IDS) can acquire information related to the attack techniques, helping in the development of prevention systems. Detection is decentralized since the IDSs are distributed on a network, installed in common nodes. The collected information and its treatment are performed in a distributed way. The advantage of Distributed Systems is more scalable and robust since it has different views of the network.

In [2], a new intrusion detection system based on cross layer Interaction between network, MAC and Physical layer. But all these systems operate in a single layer of the OSI model, or do not consider the template protection interaction and collaboration between these layers.A new intrusion detection system based on cross layer interaction between the network, MAC and physical layers is used. MAC layer uses the cross layer information from network and physical layer in order to detect possible intrusions.

To provide single cross layer IDS to several layers of OSI model instead of offering IDS for each layer. Hierarchical cluster based network topology is used in cross layer IDS. This topology divides network into several clusters and selects a cluster head node which has greatest energy reserves in the cluster.

In [10], new rule based attribute selection algorithm used for detecting intruders in WSN and also different types of attacks in WSN. On different types of attacks, it mainly focused on DOS attacks by using rule based Enhanced Multiclass SVM algorithm. The Results show that the proposed algorithm achieved high detection accuracy and reduced false alarm rate with respect to DOS attacks in WSN.The advantage is to reduce power consumption in WSN by reducing the number of packets transmitted.

III. EXISTING WORK

In [4], Two steps are required for the intrusion detection system setup. They areprofilingand anomaly/misuse detection. The steps of the profiling stage are the following: i) network logs are imported and post-processed; ii)

on the basis of processed messages, decision trees and mean and standard deviation models are produced depending on the Agent that is processing the data. Decision tree approach isused by the Central Agent, that performs a preliminary learning phase by analyzing all network traffic. Local Agents characterize monitoring parameters by estimatingthe mean and standard deviation for the normal behavior; iii) detection rules are defined according to the used technique; iv) the effectiveness of detection rules is tested through experimental data.

3.1 Sinkhole Attack

The purpose of a sinkhole attack is to gain access to alltraffic in the area of the WSN in order to launch moresevere attacks. To reach this purpose, the attacker tries toattract packets to a compromised node that belongs to thenetwork under attack. The sinkhole attack was implemented on the compromised nodes.

3.2 Sleep Deprivation Attack

The purpose of sleep deprivation attack is to hinder nodes from going into sleep mode and saving energy. The consequence is that the low energy resources available on WSN motes are soon consumed and the service offered by attacking nodes is no longer available. The sleep deprivation attack was implemented by flooding the attacked nodes with routing packets, thus forcing the receiver to process the packets and to delay the activation of the sleep mode.

On the compromised node, when routing activities are being performed and the current time is within the attack time window, a packet received is forwarded to target nodes for a high number of times. Even if that packet was discarded by the node (e.g. Because it is a duplicate), the receiver would anyway process the information before discarding it, thus consuming energy resources.

IV. PROPOSED WORK

In the proposed system, it enhanced the work of an existing system by using a decision tree technique and also it deals with Intelligent IDS methods such as Back Propagation Network,Bayesian Classification,Support Vector Machines for detecting cross layer attacks.

In the existing system, IDS acts as a second line of defense to secure the network against the attacks which is not detected by the IPS. In [2], most of the IDS concentrates on the attacks of network layers and leaves the physical, MAC and Application Layer. Most of the work mainly focuses on the stationary data.

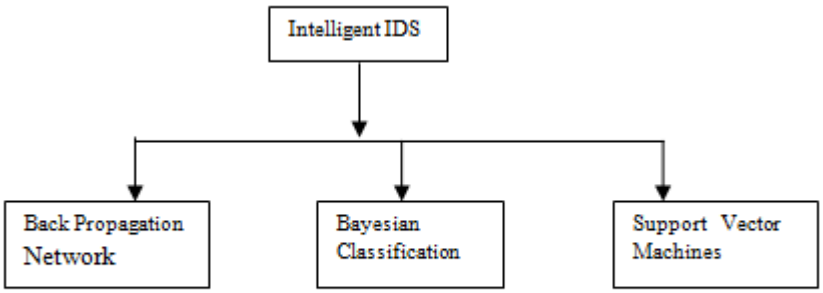


Fig 2: Types of Intelligent IDS

4.1 Decision Tree Method

In [4], Anomaly detection uses threshold metrics in local agent also called as sensor nodes and it sends alerts to the central agent if it finds any abnormalities. Misuse detection uses a decision tree method for central agent also called as a base station for detecting type of attacks. In existing [4],it concentrates only sinkhole attack and sleep deprivation attack. A Decision Tree is a flowchart like tree structure, where each internal node called as non-leaf node denotes a test on an attribute, each branch represents an outcome of the test and each leaf node called as a terminal node contains the class label. The topmost node in a tree called as root node.

4.1.1 Algorithm

Input: Dataset contains the list of attributes and split attributes by using splitting criteria

Output: A Decision tree

Method:

1:create a node K;

2:if the attributes in S are all same class, D then

3: return K as a leaf node labeled with class D;

4:ifattribute_list is empty then

5: return K as a leaf node labeled with majority class in S;

6:applyAttribute_selection_method (S,attribute_list)to findbestsplitting_criteria;

7:Label node K with splitting criteria;

8:if splitting criteria arediscrete_valuedand multiway splits allowed then

9:attribute-list<-attribute-list-splitting attribute;

10:for each result j of splitting-criteria

11:letS_j be set of data attributes in S satisfying results j;

12:ifS_j is empty then

13: attach a leaf considered with majority class in S to node K;

14:elseattach node returned by Generic_decision_tree(S_j,attribute_list) to node K;

endfor

15:return K;

Table 1

Implementation of Decision Tree using Matlab

Technique	FPR	FNR	ACC	TPR	TNR
Decision tree	**0.42**	**0.03**	**0.8**	**0.14**	**0.12**

In[4], network simulator is used for detecting attacks by using Decision Tree. In proposing System, Matlab tool is used for detecting attacks and compare the accuracy values.

Table 2

Accuracy Comparison

Attacks	Accuracy

Sinkhole	**0.7**
Sleep deprivation	**0.8**

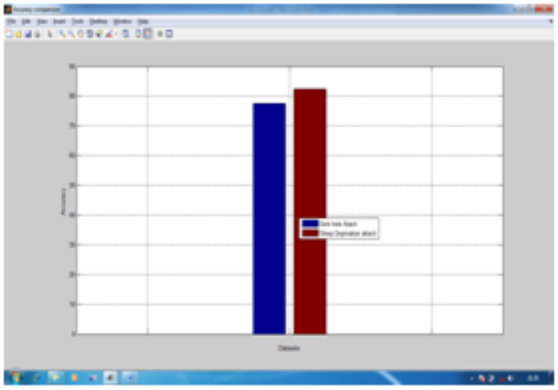


Fig 3: Accuracy comparison between sinkhole and sleep deprivation attack

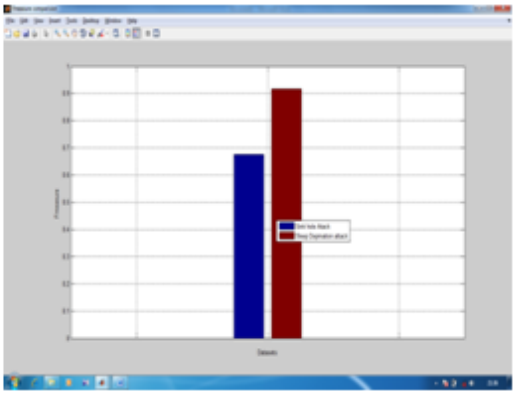


Fig 4: F-Measure Comparison

The complexity in creating large decision trees mandates people involved in preparing decision trees having advanced knowledge in quantitative and statistical analysis. This raises the chance of having to train people to complete a difficult decision tree analysis.

With reference to an existing system [4], the decision tree method is enhanced to improve accuracy better than the existing system and alsoit reduced false positive rates.

4.2 Backpropagation Network

Back Propagation is a neural network learning algorithm. A neural network is a set of connected input/output units in which each connection has a weight associated with it. Neural Network learning is also called as connectionist learning due to the connection between units. The network structure of BPN consists of three layers, as well as an input layer, a hidden layer, an output layer and many links between each layer and each layer has several processing units. The input layer is used to input the outer environmental messages, and by the intersect computing in the hidden layer, a related output is gotten from output layer.

The Back Propagation algorithm performs learning in a multilayer feed forward neural network The units in input layer called as input units. The units in the hidden layer and output layers called as output units.

Input Layer Hidden Layer Output Layer

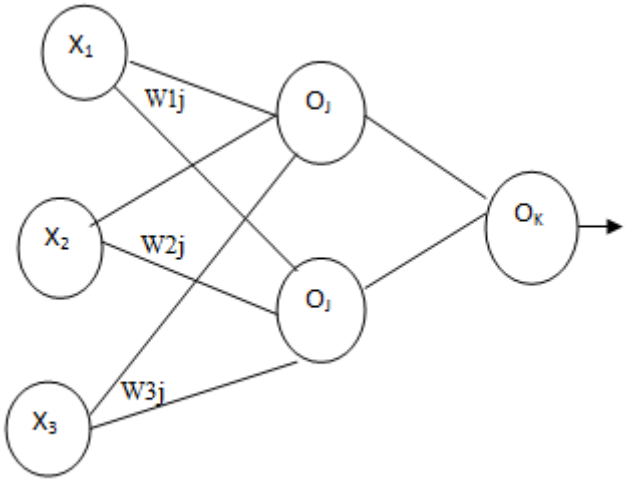


Fig 5: Back Propagation Network

The parameters such as Node id, Path, Source, Destination, Received Signal Strength can be given as input to the input layer and weight values can be combinedand send to the hidden layer. Thenit feed forwards to the output layer. It can verify whether it is normal or abnormal data and it gives results.

4.3 Bayesian Classification

Bayesian classifiers called as statistical classifiers. It can predict class membership probabilities, such as the probability that a tuple belongs to a particular class. It is based on Bayes theorem. It compares with classification algorithm and found simple Bayesian classifier called as a naïve Bayesian classifier. It compares their performance with decision tree and selected neural network classifiers. It has high accuracy and speed when it applies to large databases. The advantage of Bayesian classifiers,it has minimum error rate when compared with other classifiers.

Bayesian belief networks also called as belief networks or Bayesian networks and probabilistic networks. It can also used for classification. It contains two components such as a directed cyclic graph and set of conditional probability tables.

4.4 Support Vector Machines [SVMS]

It is a new method for classification of both linear and non- linear data. It is one of the supervised learning model with associated learning algorithms and it can analyze the data and recognize patterns and used for classification and regression analysis. It constructs hyper plane or set of hyper planes with high dimensional space used for classification, regression and other tasks.

In linear SVM, the classifier is a separating hyper-plane.In non-linear SVM, it locates a separating hyper-plane in the feature space and it classifies points in that space.The kernel function plays the role of the dot product in the feature space. The properties of SVM are Flexibility in choosing a similarity function., It has the ability to handle large feature spaces. SVM has been used effectively in many real-world problems such as text (and hypertext) categorization, Image classification, Bioinformatics(Protein classification, Cancer classification)Hand-written character recognition.

By comparing all these classification techniques and find the best techniques by using these BPN, Bayesian and support vector machine algorithms and find Cross Layer attacks in WSN to improve detection rate and accuracy.

V. CONCLUSION

In the paper, the real sensor mote data are used. The Decision tree approach is implemented for anomaly and misuse detection to improve detection rates and accuracy and reduce false positive rate. It is used to detect sink holes and sleep Deprivation attacks. With reference to an existing system, these two attacks are detected and accuracy is achieved.

REFERENCES

[1] Ana Paula R. da Silva,Marcelo H.T. Martins,Bruno P.S. Rocha, Antonio A.F. Loureiro, Linnyer B. Ruiz,HaoChiWong,"Decentralized Intrusion Detection inWireless Sensor Networks'',Q2SWinet'05 Montreal, Quebec, Canada, October 13, 2005.

[2] DjallelEddine Boubiche1 and AzeddineBilami, "Cross Layer Intrusion DetectionSystem, ForWireless Sensor Network", International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2, March 2012.

[3] Kaliyamurthie K. P. And Dr. R. M. Suresh, "Artificial Intelligence Technique Applied to Intrusion Detection", International Journal ofComputer Science and Telecommunications Volume 3, Issue 4, April 2012.

[4] Luigi Coppolino,Salvatore D'Antonio, AlessiaGarofalo, Luigi Romano"Applying Data Mining Techniques to Intrusion Detection in Wireless Sensor Networks'' Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 2013.

[5] ManasiGyanchandani, J.L.Rana, R.N.Yadav, "Taxonomy of Anomaly Based Intrusion Detection System: A Review",InternationalJournal of Scientific and Research Publications, Volume 2, Issue 12, 1 ISSN 2250-3153,December2012.

[6] Prabhakaran Kasinathan, Claudio Pastrone, Maurizio A. Spirito, Mark Vinkovits, "Denial-of-Service detection in 6LoWPAN based Internet of Things", IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), 2013.

[7] Shun-Sheng Wang, Kuo-Qin Ya,Shu-Ching Wang, Chia-Wei Liu, "An Integrated Intrusion Detection System for Cluster-based Wireless Sensor Networks'', Expert Systems with Applications 38 15234–15243, 2011.

[8] Vokorokos L., A. BalazandJ. Truelove, " Distributed Intrusion Detection System, Self Organizing Map", INES 2012-IEEE 16th International Conference on Intelligent Engineering Systems, Lisbon, Portugal-June 13–15, 2012.

[9] Yan K.Q., S.C. Wang, C.W. Liu, "A Hybrid Intrusion Detection System of Cluster based Wireless Sensor Networks", Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, Hong Kong, March 18- 20,2009.

[10] Anand, S. Ganapathy, P. Yogesh, A. Kannan, "A Rule Based Approach for Attribute Selection and Intrusion Detection in Wireless Sensor Networks",Procedia Engineering 38 1658—1664,2012.

[11] Mohammad SazzadulHoque, Md. Abdul Mukit and Md. Abu NaserBikas, "An Implementation Of Intrusion Detection System Using Genetic Algorithm",International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2, March 2012.

[12] Dewan Md. Farid1, Nouria Harbi1, and Mohammad ZahidurRahman, "Combining Naive Bayes And Decision Tree For Adaptive Intrusion Detection",International Journal of Network Security & Its Applications (IJNSA), Volume 2, Number 2, April 2010.

[13] K.V.R. Swamy, K.S. Vijaya Lakshmi, "Network Intrusion Detection Using Improved Decision Tree Algorithm",K.V.R.Swamy et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (5), 2012,4971 – 4975.

[14] ManoranjanPradhan, Sateesh Kumar Pradhan, Sudhir Kumar Sahu, "Anomaly Detection Using Artificial Neural Network",International Journal of Engineering Sciences & Emerging Technologies, April 2012. ISSN: 2231 – 6604 Volume 2, Issue 1, pp: 29-36.

[15] Hari Om &TanmoyHazra, "Statistical Techniques In Anomaly Intrusion Detection System",International Journal of Advances in Engineering & Technology, Nov. 2012. ISSN: 2231-1963.

HIGH PERFORMANCE AND LOW POWER ASYNCHRONOUS DATA SAMPLING WITH POWER GATED DOUBLE EDGE TRIGGERED FLIP-FLOP

[1]R.Ramya, [2]C.Hamsaveni

[1,2]PG Scholar, Department of ECE, Hindusthan Institute Of Technology, Coimbatore.(India)

ABSTRACT

Power consumption and energy efficiency is a major role in sequential circuit design. Power gating is a technique that is used to reduce the static power consumption of idle modules. Usage of Dual Edge Triggered Flip-flop (DETFF) is an efficient technique since it consumes the clock frequency and less power than Double Edge Triggered Flip-flops (DETFF's). Integrating power gating technique with DETFF reduces the power consumption and leakage power further, but it leads to asynchronous data sampling problem. In this paper, two methods have been used to eradicate the asynchronous data sampling problem and their power analysis has been estimated. In order to reduce the leakage power consumption further, a new design has proposed for a DETFF. Based on his new design, the two methods have been implemented using 130 μm Tanner EDA tool.

Keyword : ***Double Edge Trigger Flip Flop, Clock Gating, Power Gating, Single Edge trigger Flip Flop***.

I. INTRODUCTION

Power efficiency and energy savings are considered to be vital issues for designers. Normally, high-performance chips will have high clock frequency, which leads to high power consumption. Therefore, less power consuming designs are needed.

The major source of power consumption in a sequential circuit is clock tree and the timing components. Higher speed of clock, increase level of integration and technology scaling are reasons for high increases in power consumption. Therefore low power consumption is becoming very crucial factor for VLSI circuits. Performance assessment of the SVM showed leak size, location is both predicted with a reasonable degree of accuracy. The location prediction limits the set of locations that need to be considered when searching for a leak, thereby providing useful information for authority [1]. A set of novel D-type double edge triggered flip-flops which can be implemented with fewer transistors than any previous design. The analysis includes an implementation independent study on the effects of input sequences, in this energy dissipation of single and double edge triggered flip flops.

The system level energy savings possible by using registers consisting of double edge triggered flip flops, instead of single edge trigger flip flops [2].

The requirements of the energy dissipate high density circuits and to extend the battery life in portable systems such as devices with wireless communication capabilities. Flip Flops are mostly energy power consumed device.

A significant amount of energy is wasted to conservatively ensure power synchronization among different components [3]. A sequential circuit by a quaternary variable and uses this representation to propose and analyze two clock gating techniques. Based on it, two types of clock-gating were introduced to form a derived clock. [4]. A new simulation and optimization approach is represented, for a high performance and power issues. The analysis of an approach reveals that sources of performance and power, a set of consistent analysis approach and simulation conditions has been introduced [5].

Flip-flops use new gating techniques that reduce power dissipation to deactivating the clock signal. To overcome the presented clock duty cycle limitations of previously reported gated flip-flops. Numerical simulations of the circuit extracted from the layout with the inclusion of parasitic, show that a significant power dissipation reduction is obtained if input signal switching activity is low. [7]. The power consumption of a clock system is one of the main sources of power dissipation, typically 20 to 45% of total chip power. Consequently, many ingenious techniques have been proposed recently to reduce the clock power of the flip flops [8]. A low swing clock double-edge triggered flip flop (LSDFF) is developed to reduce power consumption significantly compared to conventional flip flops.

II. EXISTING METHOD

2.1 D Type Flip Flop

The methodologies for leakage power reduction are categorised into two classes depending on whether they reduce standby or runtime leakage. Several techniques have been proposed for standby leakage reduction. Variable threshold voltage MOS technique adjusts the device threshold voltage by body biasing. Multi threshold CMOS (MTCMOS) technique uses low voltage devices to implement main circuit elements, and high voltage devices to implement switches to disconnect the main circuit from supply line in standby mode. The proposed circuits deploy that reduced swing clock and swing data to manage dynamic power. Furthermore, it employs clock gating and power gating process during idle mode to it's eliminate dynamic power and reduce static power, while retaining its state.

The static structure of the circuit makes it feasible to be used in variable frequency power control designs. The proposed circuits were used to construct a new low-power dual-edge triggered state-retention scan FF called DET_SRSFF.

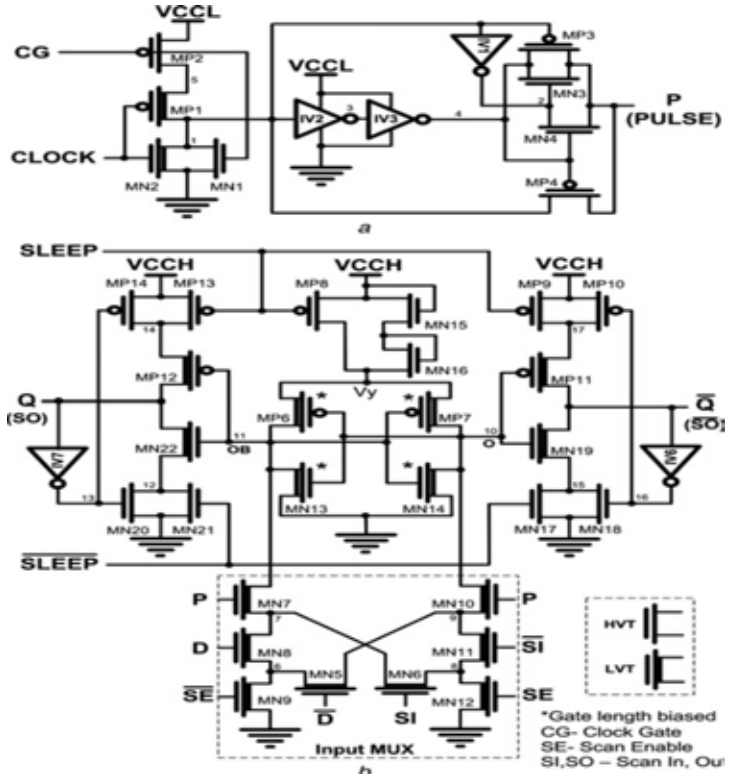


Fig 2.1 Dual-Edge Triggered State Retention Scan Flip-Flop (DET SRSFF).

The proposed FF reduces static and dynamic power consumption in both the clock tree and the FFs. For continuous operation of DET_SRSFF between the idle and active modes, a special buffer called leakage-feedback buffer is used to avoid floating output nodes, and at the same time to hold the state of the FF in the idle mode. The overall PDP of DET_SRSFF is comparable with conventional high-performance FFs and at the same time with extra level conversion and state retention feature.

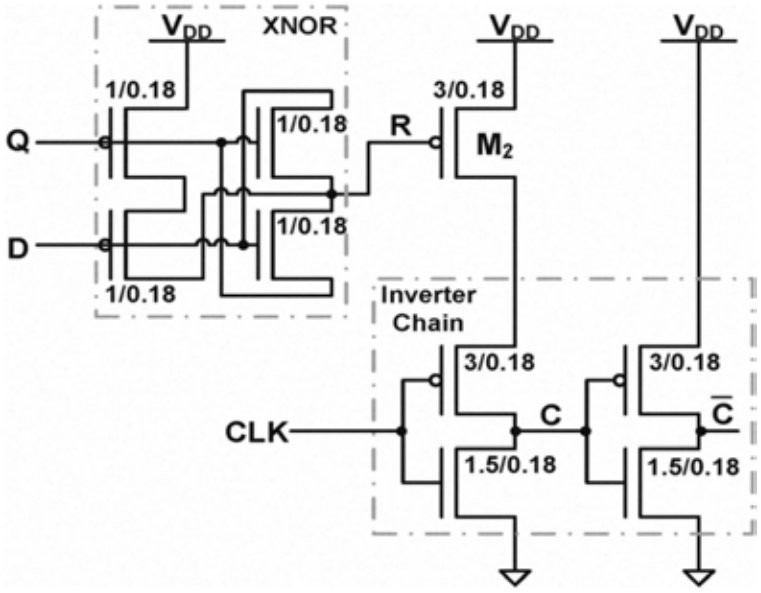


Fig 2.2 Clock-gating and internal clock generating circuitry.

2.2 Clock Gating Circuit

The proposed DHSCGFF does not require any pulse generator; it reduces the power dissipated on the clock network. The efficiency of the proposed DET-FF can be further enhanced by introducing a clock-gating circuit. This simple and energy efficient clock-gating circuit is based on

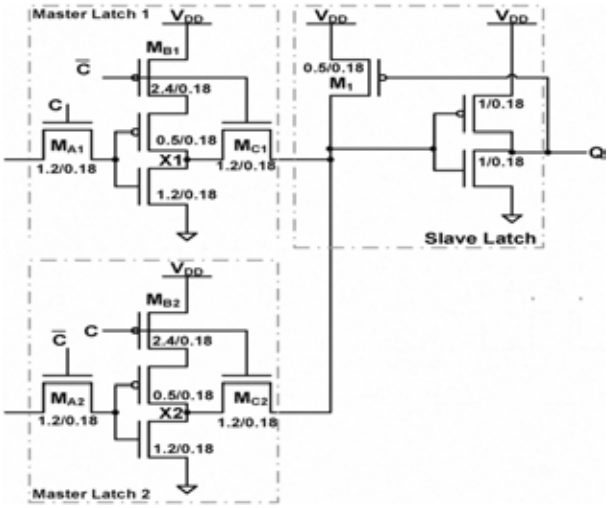


Fig 2.3 Flip-Flop circuitry. Master Latch 1 is referred to as the upper path, and Master Latch 2 is referred to as the lower path in this paper.

XNOR circuit constructed by pass transistors [8]. The pass transistor logic simplifies the circuit and reduces the internal power dissipation.

III. PROPOSED METHOD

3.1 Power Gated D Flip-Flop

The pulse generator is used which produce the dual pulse which is active at both rising and falling edge of the clock. The C (internal gated clock) signal maintains its value instead of generating an active edge in the gating mode. C changes after the transition on CLK in the non-gating mode.

Asynchronous data transition occurs in DET_SRSFF, when there is an input change while CLK equals 0. Because when there is a change in the input, clock signal is made inactive. At that time when the input is stable that means no significant change in the output. But still at that time circuit evaluate the input. This is basically used to control the discharge path. The dual triggered pulse generator produces a brief pulse signal synchronized

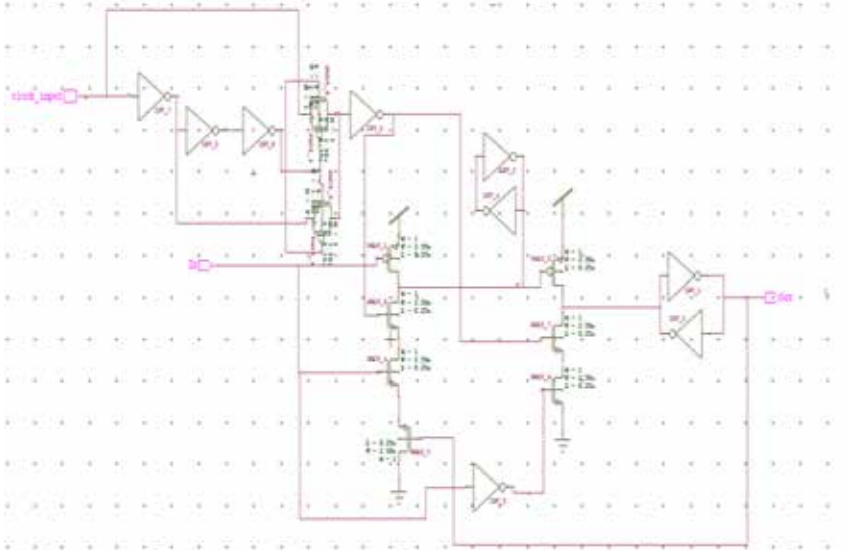


Fig 3.1 Power gated double edge triggered flip flop

Conditional precharge technique is used for removing the redundant transitions of the flip-flop to reduce the power dissipation. The schematic of this type of circuit is shown in fig 4. In this conditional technique for preventing the precharging of internal node discharging path is controlled when the input remains is high for long time. The flip-flop's output is examined and the transition is allowed only if there is a significant change in the output of the flip-flop. The correct choice of flip-flop and its corresponding design has a deep effect in reducing the power consumption. Pulse triggered flip-flops gave better output as compared master slave latch flip-flops because of timing issues.

There are different types of the dual edge triggered flip-flop used in the different synchronous circuits. There are many microprocessors which use master-slave and pulse triggered flip-flops. Master slave dual edge triggered flip flop which is made up of two stages, one is master and other is slave. They are characterized by the positive set up time and large D to Q delay. Also there is duplicating of the latch part one is for master and other is for slave. Examples of master-slave flip-flops include the transmission gated transmission gated based flip-flop, push-pull dual edge flip-flop and transmission gate latch mux (TGLM). In pulse triggered flip-flops, one is implicit pulse triggered flip-flop in which for generating the clock pulse implicit pulse generator is used and other one is explicit pulse triggered flip-flop in which generation of the clock pulse by explicit pulse generator.

V. EXPERIMENTAL RESULTS

5.1 Proposed Dual Edge Triggered Flip-Flop

In this proposed circuit of flip-flop some type of controlling circuit is embedded so that clock is disabling when the input invokes no output change. In order to eliminate the redundant transitions this data dependent technique based flip-flop is proposed. This results in saving of the power because the clock is disable at the point when no significant change at output because of stable input.

5.1.1 Result

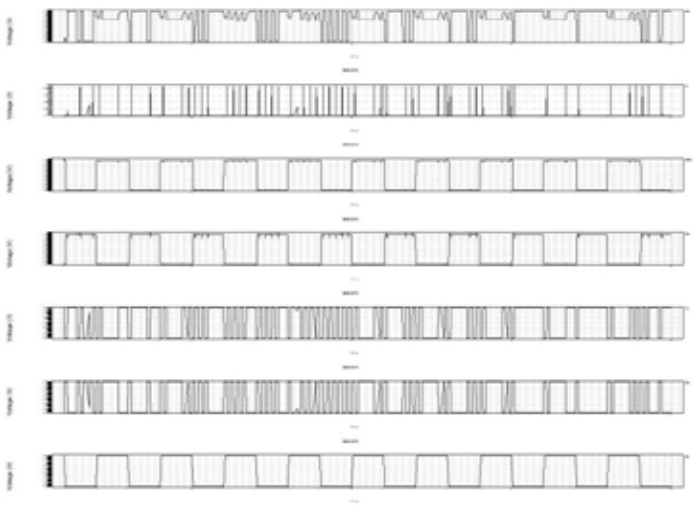


Fig 5.1 Simulation Output

5.1.2 Power Results

V1 from time 0 to 2e-006
Avg power consumed -> 1.032937e-003w
Max power 4.580270e-003 at
Time 9.29607e-008
Min power 4.637432e-009 at
Time 9.3e-007

VI. CONCLUSION

Various power reduction techniques emerged as a result of high demand in mobile devices. DETFF is an efficient technique for power reduction, when used separately. When clock gating technique is integrated with DETFF, asynchronous data sampling problem arises at the output between two clock edges. This problem has been defined in detail and solutions were given to eradicate it. Three simple approaches were made to reduce the power consumed in DETFF's by eliminating the asynchronous data sampling issue. In order to reduce the power consumption further, a new design has been proposed and based on that, three designs were implemented using Tanner EDA Tool.

REFERENCES

[1] M. Keating, D. Flynn, R. Aitken, A. Gibbons, and K. Shi, Low Power Methodology Manual: For System-on-Chip Design. New York: Springer, 2007.

[2] M. Pedram, "Power minimization in IC design: Principles and applications," ACM Trans. Design Autom. Electron. Syst. (TODAES), vol. 1, no. 1, pp. 3–56, 1996.

[3] V. Venkatachalam and M. Franz, "Power reduction techniques for microprocessor systems," ACM Comput. Surveys (CSUR), vol. 37, no. 3, and 2005.

[4] C.-C. Yu, "Low-power double edge-triggered flip-flop circuit design," in Proc. Int. Conf. Innovative Computing Information and Control (ICICIC), 2008.

[5] A. G.M. Strollo, C. Cimino, and E. Napoli, "Power dissipation in onelatch and two-latch double edge triggered flip-flops," in Proc. IEEE Int. Conf. Electronics, Circuits and Systems, 1999, vol. 3, pp. 1419–1422.

[6] R. Hossain, L. Wronski, and A. Albicki, "Low power design using double edge triggered flip-flops," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 2, no. 2, pp. 261–265, 1994.

[7] A. G. M. Strollo, E. Napoli, and C. Cimino, "Analysis of power dissipation in double edge-triggered flip-flops," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 8, no. 5, pp. 624–629, Oct. 2000.

[8] N. Nedovic and V. G. Oklobdzija, "Dual-edge triggered storage elements and clocking strategy for low-power systems," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 13, no. 5, pp. 577–590, Oct. 2005.

[9] V. Stojanovic and V. G. Oklobdzija, "Comparative analysis of master slave latches and flip-flops for high-performance and low-power systems," IEEE J. Solid-State Circuits, vol. 34, no. 4, pp. 536–548, Apr. 1999.

[10] K.-H. Cheng and Y.-H. Lin, "A dual-pulse-clock double edge triggered flip-flop for low voltage and high speed application," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS), 2003, vol. 5, pp. 425–428.

[11] K. Inoue and M. Kaneko, "Variable-duty-cycle scheduling in doubleedge- triggered flip-flop-based high-level synthesis," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS), 2011, pp. 550–553.

[12] P. Zhao, J. McNeely, P. Golconda, M. A. Bayoumi, R. A. Barcenas, and W. K. W. Kuang, "Low-power clock branch sharing double-edge triggered flip-flop," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 15, no. 3, pp. 338–345, Jun. 2007.

Biographical Notes

Ms . R . RAMYA is presently pursuing M.E. final year in Electronics & communication Engineering Department (Specialization in VLSI Design) from Hindusthan institute of technology, coimbatore, India.

Ms . C . HAMSAVENI is presently pursuing M.E. final year in Electronics & communication Engineering Department (Specialization in VLSI Design) from Hindusthan institute of technology, coimbatore, India.

ENERGY OPTIMIZATION IN MANETS USING ON-DEMAND ROUTING PROTOCOL

K.Sumathi[1], A.Priyadharshini[2]

[1, 2] *Department of CSE, Coimbatore Institute of Technology, (India)*

ABSTRACT

*Mobile Ad hoc networks (MANET) allow a set of wireless hosts to exchange information without any special infrastructure. The projct entitled "**Energy Optimization in Manets using On-Demand Routing Protocol**" motivates the need for energy management in ad hoc wireless networks. Limited battery power is one of the most important issues in mobile ad hoc network as the mobile nodes operate in limited battery power. Also there occurs a problem of broken links due to the lack of energy which cause disorder in network system. Such problem occurs due to the unawareness of energy of mobile neighbor nodes. This paper presents the implementation of Adaptive HELLO messaging scheme to determine the local link connectivity information for monitoring the link status between nodes along with the incorporation of Dynamic On Demand Routing Protocol to reduce the energy consumption of mobile nodes to certain extent.*

Keywords: DYMO, Energy Consumption, Hello Message, Mobile Ad Hoc Networks, Routing, RREQ, RRER, RREP.

I. INTRODUCTION

1.1 Introduction to Mobile Ad Hoc Network (Manet)

MANET is one of the most emerging fields in research and development of wireless network. As the popularity of mobile device and wireless networks increased significantly over the past years, it has now become one of the most vibrant and active field of communication in wireless technology.

MANET is a self configuring and infrastructure- less network. Each device or node is free to move independently, and will therefore change its links with other devices frequently in any direction. The primary challenge in creating a MANET environment is to continuously maintain the information required to route the traffic properly. Such networks can operate by themselves or by connecting itself to the larger Internet. They may contain one or more transceivers. This results in a highly dynamic and autonomous topology.

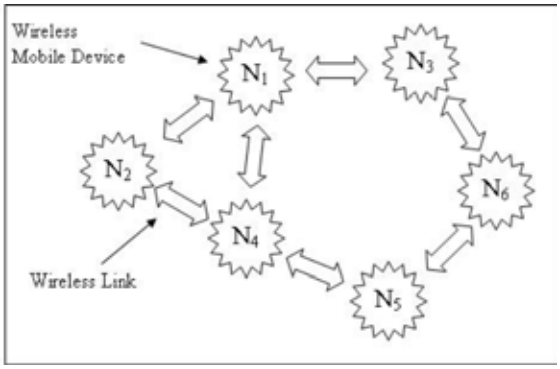


Fig.1.1 Mobile Ad Hoc Network (MANET)

MANET has routable networking environment to process the exchange of information or packet from one node to other node. Different protocols are simulated for measuring the packet drop rate, the overhead introduced by the routing protocol, end-to-end delay of packet, network throughput, etc.

This paper proposes an implementation of Adaptive Hello messaging scheme and Dynamic on-Demand routing protocol to establish a link and efficiently utilize the energy to enhance the life of network.

The rest of this paper explains the advantages of MANET, various routing protocols is MANET, overview of previous proposals, including proposed work, finally the last section contains the performance evaluation of proposed system.

1.2 Advantage of Going to Manet

Ad hoc networks are suited for the situations where an infrastructure is unavailable, and it is simple and fast, not cost effective to deploy too. The following are some of the important application related to MANET,

- Ø Business application,
- Ø Military application,
- Ø Emergency operations,
- Ø Home, office, and educational applications,
- Ø VANET,
- Ø Wireless sensor networks, mesh networks,etc.

1.3 Routing Protocol in Manet

MANET Routing Protocols are typically subdivided into two main categories: Proactive Routing Protocols and Reactive Routing Protocols.

1.3.1 Proactive Routing Protocol

Proactive routing protocol is the one in which each node maintains its route to all other network nodes. The route creation and maintenance are performed by both periodic and event-driven messages. The various proactive protocols are Destination Sequenced Distance Vector (DSDV) [1], Optimized Link State Routing (OLSR) [2].

1.3.2 Reactive Routing Protocol

In Reactive routing protocol, the route between two node is discovered only when it is demanded which is considered as the important advantage since message overhead is reduced i.e., total number of control packet transmission is reduced. There are different types of reactive routing protocols such as Ad hoc On-Demand Distance Vector (AODV) [3], Dynamic Source Routing (DSR) [4], Dynamic MANET On-Demand (DYMO) [5].

1.3.3 Hybrid Routing Protocol

An example of a hybrid routing protocol that combines both proactive and reactive approach, which brings the advantage of both the approaches together is Zone-Based Hierarchical Link-State Routing Protocol (ZRP) [6]. ZRP defines each node a zone around itself containing all neighbor nodes with certain ‘k’ hop (k=1,2 or 3). If the destination node’s position is within the zone of source then it uses proactive routing else it uses reactive

routing protocol.

1.4 Working of Routing Protocol

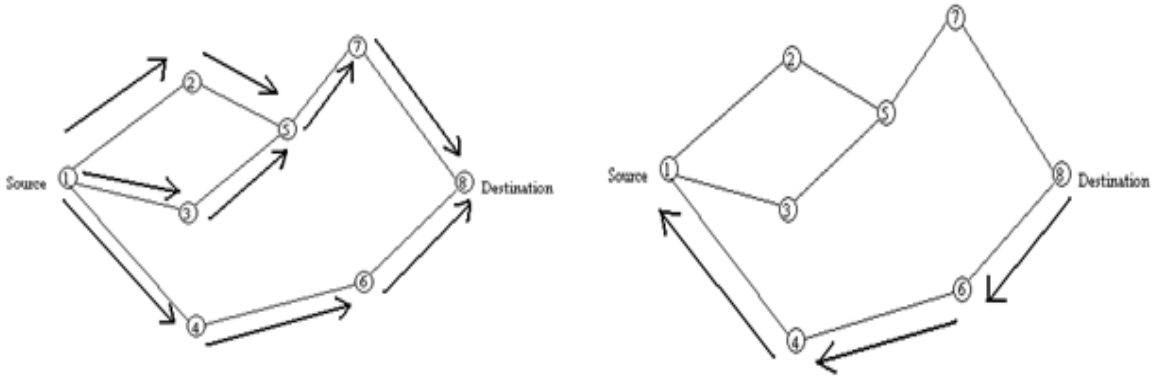
Working of the Routing Protocol consists of 2 phases: Route discovery, Route maintenance.

1.4.1 Route Discovery

When a node desires to send packets to a destination node, it first establishes a path to it for communication. The node begins the route discovery by broadcasting a route request (RREQ) message containing the IP address of the destination. When an intermediate node receives the RREQ, it records the reverse route toward the source and checks whether it has a route to the destination. If a route to the destination is not known, the intermediate node rebroadcasts the RREQ or if it has recent information about a route to the destination, route reply (RREP) message is generated. This RREP is unicast back to the source using the reverse route that is been recorded. When a RREP reaches the source, it begins to send data packets to the destination along the discovered path. If more than one RREP is received by the source, the route with the lowest hop count to reach the destination is selected.

1.4.2 Route Maintanence

This is the phase where the maintanence of link is preserved when broadcasting the packets.When a link breaks along an active path, the node upstream of the break detects the break and creates a route error (RERR) message. This message lists all destinations that are now unreachable, due to the link breakage and this information is sent to the source. Each intermediate hop deletes any broken routes and forwards the RERR packet towards the source. When the source receives this, it determines whether the packet still needs to be forwarded. If so, it begins the route discovery process for forwarding.



(a) Propogation of Route Request (RREQ) Packet (b) Path taken by the Route Reply (RREP) Packet

Fig.1.2 Propogation of RREQ, RREP

1.5 Energy Conservation in Manets

Battery energy is said to be a rare resource, and it often affects the communication activities between nodes in network. Communication takes place through direct links or through multi hop links. Due to the limited battery energy of mobile nodes, the lifetime of node becomes the key challenge. Controlling the transmission power significantly reduces the energy consumption for sending data packets and also increase lifetime of network. Nodes adjust the transmission power so as to achieve the minimum energy consumption according to the local information.

The idea of distributed power control can be used to improve energy efficiency of routing algorithm in MANET. There are some control messages such as RREP in On-Demand Routing Protocol which provide a strong indication that messages should trigger a node to switch to active node from sleep. Since the communication with a neighbor is only possible if the neighbor is in active mode, it is necessary for nodes to track energy modes of neighbors i.e., active, sleep or idle. The neighbor's power mode can be discovered in two ways: the first way is through explicit local HELLO message exchanges with piggybacked information about the energy management mode of a node, and the second way is via passive inference.

Energy efficiency is measured by the duration of the time over which a network can maintain a certain performance level, which is usually called as the nework lifetime.Using the power consumption is not only a single criterion for conserving energy efficiency. Hence, routing to maximize the lifetime of the network is different from minimum energy routing. Minimum energy routing sometimes attract more flows since the nodes in these route exhaust their enegy very soon. Hence, the whole network cannot perform many task due to the failure of these nodes. Routing with maximum lifetime balances all the routes and nodes globally so that it can maintain certain performance level for a longer time.

Hence saving energy at the time of broadcasting in order to recover from the node failure and during re-routing around failed node is essential.

1.6 Energy Efficient Routing Protocol

Energy is said to be a limiting factor in case of ad hoc networks [7] [8] [9]. Routing in ad hoc network has some unique characteristics:

- Ø Energy of node is crucial and it depends upon battery which has limited power supply.
- Ø Nodes can move in an uncontrolled manner, so frequent route failures are possible.
- Ø Wireless channels have lower and more variable bandwidth compared to wired network.

Energy efficient routing protocols are the only solution to above situation. Most of the work of making protocols energy efficient has been done by reactive routing rather than proactive routing protocols. Energy efficiency can also be achieved by sensible flooding at the route discovery process in reactive protocols. And it can also be achieved by using efficient metric for route selection such as cost function, node energy, battery level etc.

Here the efficiency not only refers to the successful delivery of packets with less consumption of power, but also refers to the increase in duration of maintaining the link between the nodes to ensure increase in performance. This can be achieved by using AODV & DYMO routing protocol.

Consider an example, of a multi-hop communication channels from A to B through an intermediate node C between them, see Fig.1.3

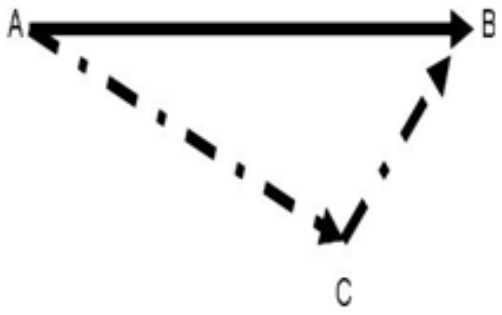


Fig. 1.3: A Simple Scenario for Energy Consumption in Multi-hop Networks

There are two possible ways to communicate between A and B. One is to directly transmit data from A to B; or relay with Node B. However, these two different methods lead to two different level of energy consumption, in which one must be better. Transmitting a packet from A to B consumes less energy rather than sending the same packet toB through C.

II. BACKGROUND

2.1 A New Method for Restoration Broken Links in Wireless Ad Hoc Networks by Estimation Energy Consumption

A novel method on energy estimation to restore broken links and reconstruction of the paths is proposed to investigate the effect of broken links on topology control and routing process in ad hoc network. It is indicated that these effects were harmful in the mentioned couple of network portions. This work [10] is used as Hardware Method for estimating energy in ad hoc node, and this method has high speed too. Hence, it helps to find out or investigate the effect of link breakage in ad hoc network. One may find that the use of routing protocol algorithm to estimate the energy and also to reconstruct the path will conflict. For this purpose a strategy was made in order to prevent link break and disordering which provide some suggestions to route the network through prediction and time estimation of link break.

2.2 Improving the Network Lifetime of Manets through Cooperative Mac Protocol Design

To relay the overhearing information to achieve greater efficiency a cooperative communication is proposed to utilize nearby terminals to transmit in wireless networks. In order to deal with the complicated medium access interactions induced by relaying and leverage the benefits of such cooperation an efficient Cooperative Medium Access Control (CMAC) protocol is needed. This paper [11] proposes a novel cross-layer Distributed Energy-adaptive Location-based CMAC protocol, namely DEL-CMAC. The design objective of DEL-CMAC is to improve the performance of the MANET in terms of network lifetime and energy. A practical energy consumption model is described that the nodes consume energy on both transceiver circuit and transmit amplifier. A distributed utility based best relay selection strategy is also incorporated, which selects the best path to route based on location information and residual energy. Also, to enhance the spatial reuse for the mobile nodes, an innovative network allocation vector setting is provided to deal with the varying transmitting power of the source and relay terminals. The result of proposed DEL-CMAC prolongs the network lifetime even for high circuitry energy consumption.

2.3 Power Management for Wireless Data Transmission Using Complex Event Processing

Energy consumption of wireless data transmission is a significant part in wireless mobile devices. It is context-dependent, i.e. it depends on internal and external contexts, such as application workload and wireless signal strength. This paper [12] proposed an event-driven framework for efficient power management on mobile devices when communicating. This framework adapts the behavior of a device component or an application to that change in contexts, according to event-condition-action (ECA) rules specified by the developer to describe the power management mechanism. It also supports complex event processing by correlating various events, and helps to discover complex events that are relevant to power consumption. This framework is evaluated with two applications in which the data transmission is adapted to traffic patterns and wireless link quality. These applications can roughly save 12 percent more energy compared to normal operation. But this paper doesnot mention about the collaborative power management between the mobile nodes.

2.4 An Adaptive Hello Messaging Scheme for Neighbor Discovery in On-Demand Manet Routing Protocols

In mobile ad hoc networks, local link connectivity information is extremely important for route establishment and maintenance. Periodic Hello messaging scheme is a widely-used to obtain local link connectivity information. However, unnecessary hello messaging can drain their battery power when mobile devices are not in use. This paper [13] proposes a messaging scheme to suppress unnecessary Hello messages without reduced detectability of broken links. Simulation result shows that the proposed scheme reduces energy consumption and network overhead without any explicit difference in throughput.

This scheme dynamically adjusts Hello intervals, and does not increase the risk of sending the packet through a broken link. To estimate unavailability of link between nodes, average time gap between two consecutive events is found. By monitoring the event intervals, state of a node can be estimated. If a node is not involved in any communication for a given period, then it does not need to maintain the status of the link, and broadcasting of hello packets during this period are unnecessary. If a constant Hello interval is used, the risk of attempting to transmit a packet through a broken link decreases as the event interval increases. Instead of using a constant Hello interval, proposed scheme uses a constant risk level. As the event interval increases, the Hello interval can also be increased without increasing risk.

This proposed scheme result in reduced battery power through practical suppression of unnecessary Hello messaging, which decrease network overhead and hidden energy consumption.

2.5 Implementation of DYMO Routing Protocol

DYMO routing protocol is proposed by Perkins & Chakeres [5]. It is defined as the successor of AODV or ADOVv2 and it keeps on updating. DYMO operates similar to its predecessor i.e. AODV and does not add any extra modifications to the existing functionality. It is a purely reactive protocol in which routes are computed on demand. Unlike AODV, DYMO does not support unnecessary HELLO message transmission; also the operation is purely based on sequence number that is assigned to all the packets. It employs sequence numbers to ensure loop freedom. It also enables on demand, multi-hop unicast routing among the nodes in a network. The basic operations are route discovery and maintenance to obtain a valid path and also to avoid the existing obliterated routes from the routing table to reduce the packet dropping in case of any route break or node failure. The simulation has been performed with varying pause times and observed that DYMO being the successor of AODV performs better in all the terms.

III. SYSTEM DESIGN AND IMPLEMENTATION

In the proposed system, the behaviour of packet switching network is simulated using opnet modeler 14.5. This network may consist of four peripheral nodes to generate traffic while a central hub node that relays the traffic to the appropriate destination within the network. The performance of the network is measured by the end-to-end delay experienced by traffic on the network when delivering the packet. This proposed system consists of following modules,

- Creation of nodes,
- Defining the packet model, and link model,
- Creating the hub model,
- Creation of peripheral node model,
- Propogation of HELLO packets

3.1 Creation of Node

With the advances in software and hardware architecture, mobile nodes can be created according to the requirements. Each node is equipped with a transmitter and receiver and they are said to be purpose-specific, autonomous and dynamic. This characteristic compares greatly with fixed wireless network as there is no master slave relationship that exists in ad hoc network. Mobile nodes rely on each other to the established communication, thus each node acts as router. Therefore, packets can travel directly or through some set of intermediate nodes to reach destination node.

3.2 Defining The Packet Model And Link Model

The packets in this network contain a single field associated with the destination address in it. After the packet format has been created, it is specified as an attribute in a generator so that it can be formatted accordingly. The packet contains attributes such as name, type, size. Also, the set at creation attribute is changed to unset which may ensure that the field will not be assigned a default value when the packet is created.

Point-to-point links can be simplex (unidirectional) or duplex (bidirectional). Here in the proposed model, custom duplex link is used to connect transmitter-receiver pairs for end to end delivery. Link model is designed in such a way that it connects the hub and peripheral nodes and supports in packet transfer.

3.3 Creation Of Hub Model

The hub node model consists of point-to-point transceivers for each peripheral node, and a process model to relay packets from receiver to the appropriate transmitter. The packet streams have a unique index which is an easiest method to set up a direct association between the hub process outgoing packet stream indices and the peripheral destination address values. In a more adaptive model, the hub process model is made to maintain a table for translating destination address value to transmitter stream indices. In this proposed system, a direct correspondence is made between designating addresses and packet stream indices.

Also, since each packet is associated with an interrupt, the hub process model may receive an interrupt whenever a packet arrives from a receiver. In the process model, the PK_ARRVL function compares the interrupt type of arrived packet with the predefined constant OPC_INTRPT_STRM, which is an stream interrupt. This type of interrupt for this model is good to safeguard against run-time transition errors.

The macro definition for PK_ARRVL function is:

#define

(
PK_ARRVL(OP_INTRPT_TYPE()==OPC_INTRPT_STRM)
)

3.4 Creation of Peripheral Node Model

The peripheral node model generates packet, assigns destination address, and processes received packets. It uses a user-defined process model to assign destination addresses to the generated packet and transmit them to the node's point-to-point transmitter. This process model retrieves the packet arriving from the point-to-point receiver and processes it to calculate the packet's end-to-end delay and the value is written to a global statistic so that it is accessible to multiple processes throughout the system.

In order to assign the destination address and calculation of end-to-end delay, the peripheral node process model needs two states:

- Ø an initial state
- Ø an idle state

The initial state is set to load the process model with a uniform PDF in a range of 0 to 3.

address_dist = op_dist_load ("uniform_int", 0, 3);

3.5 Propogation of Hello Packet

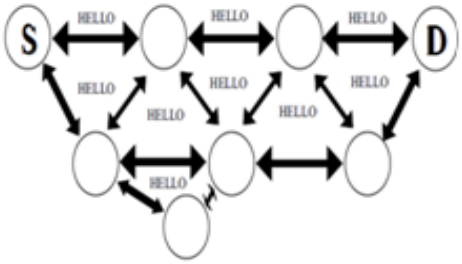


Fig.3.1: Propogation of HELLO packets

To maintain connectivity, if a node has not sent any broadcast control message within a specified interval, a hello message is locally broadcasted to the every node in the network. For better result leads at least one hello message must be transmitted for every time period. If the node is unable to receive any hello message from a neighbor for several time intervals, it indicates that neighbor is no longer within transmission range, and connectivity has been lost. Two variables are responsible for the determination of connectivity using hello messages: Hello Interval and Allowed Hello Loss. Hello Interval- specifies the maximum time interval between the transmissions of hello messages. Allowed Hello Loss- specifies the maximum number of periods of hello interval to wait without receiving a hello message before detecting a loss of connectivity to a neighbor. The recommended value for HELLO INTERVAL is one second and for ALLOWED HELLO LOSS is two.

3.6 System Requirements

Processor	**:** INTEL PENTIUM III, 4 or compatible (500 MHZ or Better)
RAM	**:** 256 MB, 512 MB
Disk space	**:** 400 MB
Display	**:**1024×768 or higher resolution, 256or more colors
Platform	**:**Windows 2000,xp,vista
Application Development	**:**Opnet modeler 14.5

OPNET stood for Optimized Network Engineering Tool. OPNET Modeler is a software tool for computer

network modelling and simulation. Through this networking software, R&D process is accelerated resulting to easy analysis and design of communication networks, application, protocols, and devices. It offers the fastest discreet event simulation engine when compared with other networking solutions in the industry. It has an open interface for easy integration of libraries, external object files and other network simulators.

IV. PERFORMANCE EVALUATION

To examine the effect of the packet generation rate in an simulation environment for measuring the performance of the network, two simulation runs with different packet interarrival times are to be considered. End to end delay and link utilization is measured. In this simulation, the source node creates packets of a constant size. This setting, in combination with the fixed data rate at the point-to-point transmitters and receivers, may result in a fixed end-to-end delay for the packets.

However, if packets are sent more quickly to a transmitter, some of the packets will be delayed in the transmitter's queue. So, if the packet interarrival time is varied, the end-to-end delay will be varied and will be affected. To model this, configure two simulation runs with different packet interarrival times.

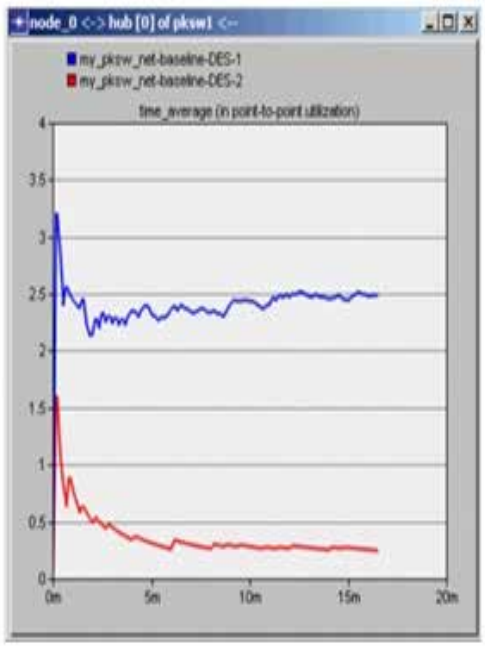


Fig.4.1: The time-averaged graph of energy utilization

End to end delay is calculated, when packets are transmitted and received and the result is shown below in linear graph.

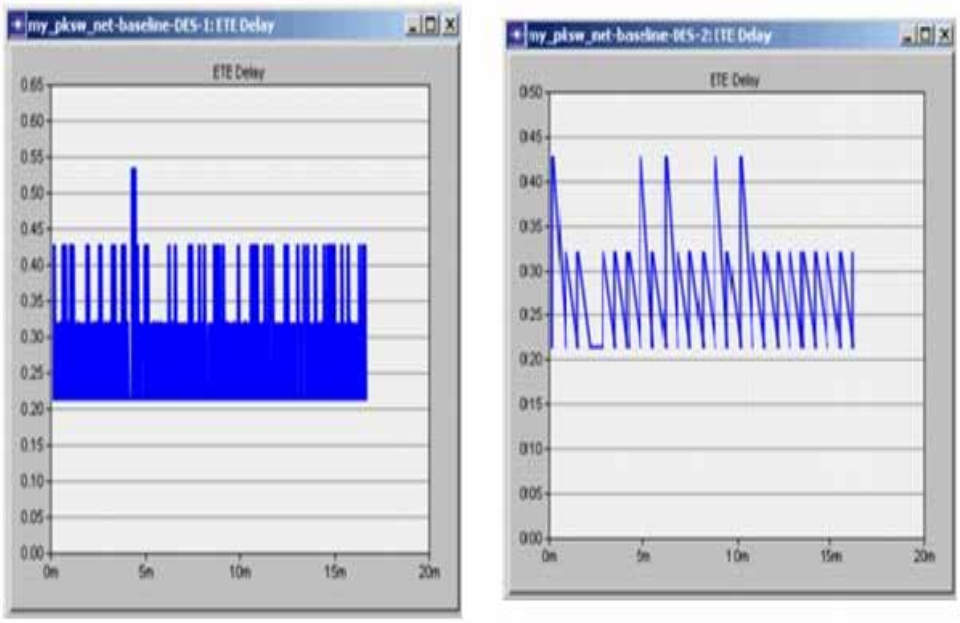


Fig.4.2: Linear graph displaying End to End delay

The above linear graph doesnot show the end to end delay clearly. So, discrete graph is used for this is shown in fig.4.3

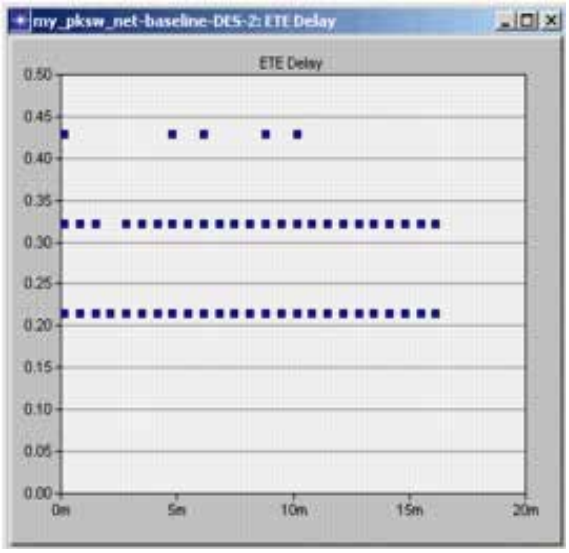


Fig.4.3: Discrete graph showing the End to End delay

V. CONCLUSION AND FUTURE WORK

A mobile ad hoc network (MANET) consists of one or more autonomous mobile nodes, each of which communicates directly or indirectly with the neighbor nodes within its radio range. The field of MANET is rapidly growing due to varied advantage and applications. Energy efficiency is a challenge faced especially in designing a routing protocol. In the proposed system, packet switching networks is implemented for transmitting and receiving the packets, finally calculate the delay.

A single routing protocol is hard to satisfy all requirements. i.e., one routing protocol cannot be a solution for all energy efficient protocol that designed to provide the maximum possible requirements, according to certain

required scenarios.

In future, Ant Routing Protocol can be used to find the optimal path, also Efficient Energy Aware Routing Protocol (EEARP) can be proposed to increase the network lifetime of MANET. AlsoUsing a mini-max formulation, EEARP selects the path that has the largest packet capacity at the smallest residual packet transmission capacity. The proposed scheme may reduce the energy consumption to some extent and decreases the mean delay, while achieving a good packet delivery ratio.

REFERENCES

[1] C. E. Perkins and P. Bhagwat, *Highly dynamic destination sequenced distance vector routing (DSDV) for mobile computers*, Proceedings of ACM SIGCOMM 94, 1994, pp. 34–244.

[2] P. Jacquet, P. Muhlethaler, T. Clausen, A. Laouiti, A. Qayyum, and L. Viennot, *Optimized Link State Routing Protocol for Ad-Hoc Networks*, in Proceedings of the 5th IEEE Multi Topic Conference (INMIC 2001), 2001, pp. 62-68.

[3] C. Perkins, E. B. Royer, S. Das, *Ad hoc On-Demand Distance Vector (AODV) Routing - Internet Draft*, RFC 3561, IETF Network Working Group, July 2003.

[4] D. B. Johnson, D. A. Maltz, Y.C. Hu, *The Dynamic Source RoutingProtocol for Mobile Ad Hoc Networks (DSR)*, IETF Draft, April 2003.

[5] I. Chakeres and C. Perkins, *Dynamic MANET On-demand (DYMO) Routing draft-ietf-manetdymo-17*, Internet Engineering Task Force, Mar. 2009.

[6] Pearlman, Marc R., Haas, Zygmunt J. *Determining the Optimal Configuration for the Zone Routing Protocol*, IEEE Journal on Selected Areas in Communications, vol. 17, Nov.1999, pp. 1395-14141.

[7] V.Kauadia and P.R.Kumar, *Power Control and clustering in ad hoc networks*, IEEE INFOCOM 2003.

[8] Juan A. Sanchez and Pedro M. Ruiz, *LEMA: Localized Energy-Efficient Multicast Algorithm based on Geographic Routing*, in Proceedings. Of 2006 31st IEEE Conference on Local Computer Networks in 2006.

[9] Dahai Du and Huagang Xiong, *A Location aided Energy-Efficient Routing Protocol for Ad-hoc Networks*, in wireless and optical communications conference (WOCC), 2010

[10] Peyman Arebi, *A New Method for Restoration Broken Links in Wireless Ad-hoc Networks by Estimation Energy Consumption*, Fourth International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN), 2012 pp. 377-381.

[11] Xiaoyan Wang, Jie Li, *Improving the Network Lifetime of MANETs Through Cooperative MAC Protocol Design*, Student Member, IEEE, and, Senior Member, IEEE,2013

[12] Yu Xiao, Wei Li, Matti Siekkinen, Petri Savolainen, Antti Yla¨ -Jaaski, and Pan Hui, *Power Management for Wireless Data Transmission Using Complex Event Processing*, IEEE Transactions on computers, vol. 61, Dec.2012,pp no.12.

[13] Seon Yeong Han, and Dongman Lee, *An Adaptive Hello Messaging Scheme for Neighbor Discovery in On-Demand MANET Routing Prototocol*, IEEE vol. 17, May 2013, no. 5.

[14] Floriano De Rango, Francesca Guerriero and Peppino Fazio, *Link-Stability and Energy Aware Routing Protocol in Distributed Wireless Networks*, IEEE Transactions On Parallel And Distributed Systems, vol. 23, Apr. 2012, No. 4.

[15] Incheon Park, Jinguk Kim, Ida Pu. *Blocking Expanding Ring Search Algorithm for Efficient Energy Consumption in Mobile Ad Hoc Network*.

[16] T. Clausen, C. Dearlove, and J. Dean, *Mobile ad hoc network (MANET) neighborhood discovery protocol (NHDP)*, 2010.

A COMPARISON BETWEEN DIFFERENT TYPES OF SOFTWARE DEVELOPMENT LIFE CYCLE MODELS IN SOFTWARE ENGINEERING

Mr. Ashish Kumar Gupta

Assistant Professor, Dept. of C.S.E., I.T.S. Engineering College, Greater Noida, U.P., (India)

ABSTRACT

This research deals with a vital and important issue in computer world. At present the software systems can not be built with mathematical or physical certainty so they are not perfect upto the level. It is concerned with the software management processes that examine the area of software development through the development models, which are known as software development life cycle. In this research paper the comparison between various software development models has been carried out. Each and every model of SDLC has its own advantages and limitations so in this research we have to describe each model on behalf of various important features. Various SDLC models like waterfall, iterative, prototype model and spiral model were suggested.

Keywords: Software Development Life Cycle, Phase of SDLC Models, Software Development Process, Comparison, Four Models.

I INTRODUCTION

SDLC stands for Software Development Life Cycle. A Software Development Life Cycle is essentially a series of steps, or phases, that provide a model for the development and lifecycle management of an application or piece of software.

Software engineering is the study and an application of engineering to the design, development, and maintenance of software. Typical formal definitions of software engineering are: "the application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software". It is the application of engineering to software because it integrates significant mathematics, computer science and practices whose origins are in Engineering.

There are various processes and methodologies have been developed over the last few decades to improve software quality, with varying degrees of success. However, it is widely agreed that no single approach that will prevent project failures in all cases.

A software development process is a structure imposed on the development of a software product. There are several models for such processes, each describing approaches to a variety of tasks or activities that take place during the process. Software Engineering processes are composed of many activities, notably the following:

- Requirements Analysis
- Specification
- Software architecture
- Implementation
- Testing
- Documentation
- Training and Support
- Maintenance

1.1 General Software Process Models are

Even though there are number of models each software Development Company adopts the best-suited model, which facilitates the software development process and boosts the productivity of its team members.

- Waterfall model: Separate and distinct phases of specification and development.
- Prototype model.
- Rapid application development model (RAD).
- Evolutionary development: specification, development and validation are interleaved.
- Incremental model.
- Iterative model.
- Spiral model.
- Component-based software engineering : The system is assembled from existing components.

There are many variants of these models e.g. formal development where a waterfall-like process is used, but the specification is formal that is refined through several stages to an implementable design.

II DIFFERENT PHASES OF SDLC

Software life cycle models describe phases of the software cycle and the order in which those phases are executed. Each phase produces deliverables required by the next phase in the life cycle. Requirements are translated into design. Code is produced according to the design which is called development phase. After coding and development the testing verifies the deliverable of the implementation phase against requirements. The basic activities or phases to be performed for developing software system are-

- Determination of System's Requirements
- Design of system
- Development (coding) of software
- System Testing

Fig1. Software Development Life Cycle

III SOFTWARE DEVELOPMENT MODELS

3.1 Waterfall Model

The waterfall model is a sequential design process, used in software development processes, in which progress is seen as flowing steadily downwards (like a waterfall) through the phases of Conception, Initiation, Analysis, Design, Construction, Testing, Production / Implementation and Maintenance.

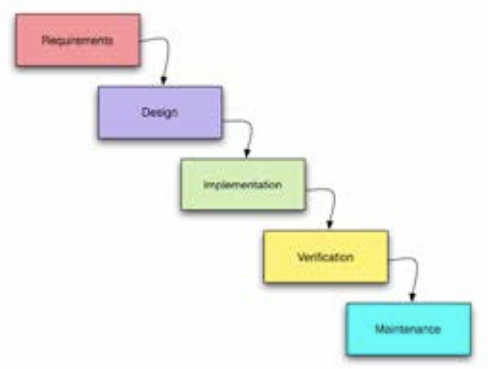


Fig2. Waterfall Model

1. **Basic Principles**
   - Problems can be solved more easily if they are more clearly defined.
   - Large amounts of code are more traceable if they are structured.
   - A good project life-cycle plan improves the development process.
   - System documentation is a byproduct of the development process, and is not done later, as an afterthought.
2. **Advantages of Waterfall Model**
   - This model is simple and easy to understand and use.
   - It is easy to manage due to the rigidity of the model – each phase has specific deliverables and a review process.
   - In this model phases are processed and completed one at a time. Phases do not overlap.
   - Waterfall model works well for smaller projects where requirements are very well understood.

3. Disadvantages of Waterfall Model

- Once an application is in the testing stage, it is very difficult to go back and change something that was not well-thought out in the concept stage.
- No working software is produced until late during the life cycle.
- High amounts of risk and uncertainty.
- Not a good model for complex and object-oriented projects.
- Poor model for long and ongoing projects.

3.2 Iterative Model

An iterative life cycle model does not attempt to start with a full specification of requirements. Instead, development begins by specifying and implementing just part of the software, which can then be reviewed in order to identify further requirements. This process is then repeated, producing a new version of the software for each cycle of the model.

1. Basic Principles

- Manage requirements not tasks, based on use cases and nonfunctional requirements.
- Manage to meet business goals, due dates and budgets. Be willing to change requirements to fit these, not the other way around.

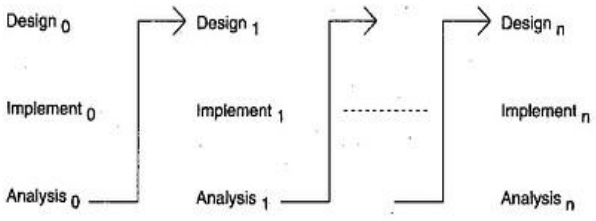


Fig3. Iterative Model

- Begin with a simple implementation of a subset of the requirements that demonstrates key aspects of the system.
- Design around isolated, easy-to-find modules that group small sets of related requirements. Complete or re-code one module per iteration.
- Work in short cycles (1-6 weeks) composed of overlapping phases: requirements, design, programming, testing.
- During the iteration, the external customer or project manager cannot change the scope for that iteration, but the development team may change the scope by dropping features if the end date will not be met.
- Any difficulty in design, coding and testing a module should signal the need for redesign or re-coding.
- Modifications should become easier to make as the iterations progress. If not, redesign is needed.

2. Advantages of Iterative Model

- It is much better model of the software process.
- It allows feedback to proceeding stages.
- It can be used for project wherein the requirements are not well understood.

3. Disadvantages of Iterative Model

- Each phase of an iteration is rigid with no overlaps
- Costly system architecture or design issues may arise because not all requirements are gathered up front for the entire lifecycle.
- No clear milestones in the development process.

3.3 Prototyping Model

The basic idea here is that instead of freezing the requirements before a design or coding can proceed, a throwaway prototype is built to understand the requirements. This prototype is developed based on the currently known requirements. By using this prototype, the client can get an "actual feel" of the system, since the interactions with prototype can enable the client to better understand the requirements of the desired system.

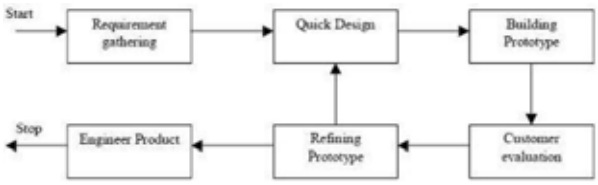


Fig4. Prototyping Model

1. Basic Principles

- Prototype model should be used when the desired system needs to have a lot of interaction with the end users.
- Not a standalone, complete development methodology, but rather an approach to handling selected parts of a larger, more traditional development methodology.
- Attempts to reduce inherent project risk by breaking a project into smaller segments and providing more ease-of-change during the development process.
- Prototyping ensures that the end users constantly work with the system and provide a feedback which is incorporated in the prototype to result in a useable system.

2. Advantages of Prototyping Model

- Users are actively involved in the development.
- Since in this methodology a working model of the system is provided, the users get a better understanding of the system being developed.
- Errors can be detected much earlier.
- Confusing or difficult functions can be identified.

3. Disadvantages of Prototyping Model

- Possibility of causing systems to be left unfinished.
- Possibility of implementing systems before they are ready.

3.4 Spiral Model

The spiral model is a risk-driven processmodel generator for software projects. Based on the unique risk patterns of a given project, the spiral model guides a team to adopt elements of one or more process models, such as incremental, waterfall, or evolutionary prototyping.

1. Basic Principles

- Focus is on cost and risk assessment throughout the life cycle.
- Useful for Long-term project commitment unwise because of potential changes to economic priorities.
- Users are unsure of their needs.
- Requirements are complex.

2. Advantages of Spiral Model

- High amount of risk analysis hence, avoidance of Risk is enhanced.

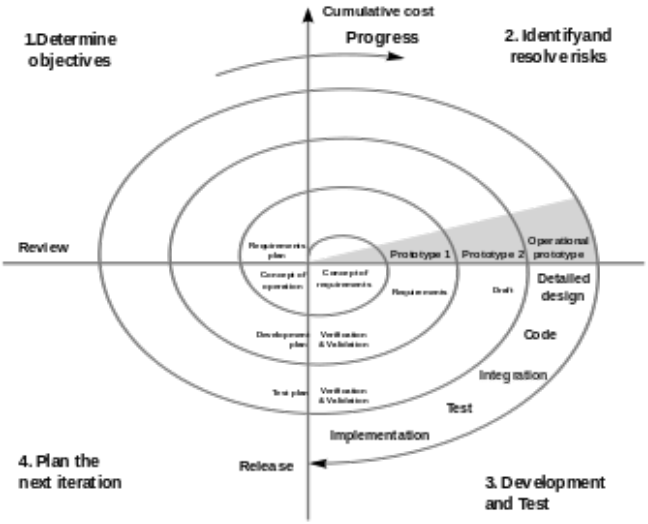


Fig5. Spiral Model

- Good for large and mission-critical projects.
- Strong approval and documentation control.
- Additional Functionality can be added at a later date.
- Software is produced early in the software life cycle.

3. Disadvantages of Spiral Model

- Can be a costly model to use.
- Risk analysis requires highly specific expertise.
- Project's success is highly dependent on the risk analysis phase.
- Doesn't work well for smaller projects.

4. Comparative Analysis of SDLC Models

Features	Waterfall Model	Iterative Model	Prototyping	Spiral Model

			Model	
Requirements Specification	Beginning	Beginning	Frequently Changed	Beginning
Cost	Low	Low	High	Expensive
Simplicity	Simple	Intermediate	Simple	Intermediate
Expertise Required	High	High	Medium	High
Risk Involvement	High	Easily Manage	Low	Low
Overlapping Phases	No	No	Yes	Yes
Flexibility	Rigid	Less Flexible	Highly Flexible	Flexible
Maintenance	Least Glamorous	Promoted Maintainability	Routine Maintenance	Typical
Reusability	Limited	Yes	Weak	High
Documentation Required	Vital	Yes	Weak	Yes
User Involvement	Only At Beginning	Intermediate	High	High
Cost Control	Yes	No	No	Yes
Resource Control	Yes	Yes	No	Yes
Guarantee of success	Less	High	Good	High

Table1. Comparison of SDLC Models

V CONCLUSION

SDLC models are tools that allow the development team to correctly follow the SDLC steps to create software that meets a business need. Each SDLC model has evolved as a new technology and has addressed weaknesses of older models. After analysis of all models through the various factors in this research, it is concluded that:

1. There are many existing SDLC models for developing systems with different requirements.
2. Waterfall model is used by various big companies for their internal projects.
3. Most commonly used models for developing systems are waterfall model and spiral model.
4. Each model has advantages and disadvantages for developing the systems, so each new model tries to eliminate the disadvantages of previous model.

REFERENCES

[1]. Ms. Shikha Maheshwari, Prof. Dinesh Ch. Jain, “A Comparative Analysis of Different types of Models in Software Development Life Cycle”, ISSN: 2277 128X (online), vol.2 Issue 5, may 2012.

[2]. Nabil Mohammed Ali Munassar1 and A. Govardhan ”A Comparison Between Five Models Of Software Engineering” International Journal of Computer Science Issues, Vol. 7, Issue 5, September 2010.

[3]. Ashish B.S..., Dr.Vinay C.., “Survey of Software Life Cycle Models by Various Documented Standards”IJCST Vol. 2, Iss ue 4, Oct . - Dec. 2011.

[4]. Chan, D.K.C. Leung, K.R.P.H, "Software development as a workflow process", 2-5 Page(s): 282 – 291, Dec. 1997.

[5]. Sanjana Taya "Comparative Analysis of Software Development Life" ISSN:2229-4333(print),ISSN:0976-8491(online),vol.2ISSUE 4,Oct-Dec2011.

[6]. Dr. Deepshikha Jamwal "Analysis of software Development Models", ISSN: 2229-4333(print),ISSN:09768491 (online), vol.1ISSUE 2, Decemder 2010.

[7]. Maglyas, A.; Nikula, U.; Smolander, K.,"Comparison of two models of success prediction in software development projects", Software Engineering Conference (CEE-SECR), 2010 6th Central and Eastern European on 13-15 Oct. 2010, pp. 43-49.

[8]. Swapanja Ingale "Comparative Study Of Software Development Model" International conference on advances in computing &management 2012.

[9]. Rothi, J.,Yen, D, "System Analysis and Design in End User Developed Applications", Journal of Information Systems Education, 1989.

[10]. S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[11]. K. Schwalbe(2009), "Information Technology Project Management", 6th Edition, Cengage Learning.