



SHANLAX

INTERNATIONAL JOURNAL OF ARTS, SCIENCE AND HUMANITIES

(A Peer-Reviewed-Refereed/Scholarly Quarterly Journal Globally Indexed with Impact Factor)

Vol. 5

Special Issue 2

January, 2018

Impact Factor: 2.114

ISSN: 2321-788X

UGC Approval No: 43960

International Conference

On Recent Trends in IT NCCSA'18

28th & 29th January 2018

Organized by

PG DEPARTMENT OF COMPUTER SCIENCE AND APPLICATIONS



ISLAMIYAH WOMEN'S ARTS AND SCIENCE COLLEGE

Recognized by UGC under section 2(f) and 12(b) of the UGC Act 1956 &
Permanent Affiliation with Thiruvalluvar University

#10, Bye-Pass Road, Newtown
Vaniyambadi-635752

CONFERENCE ORGANIZING COMMITTEE

Chief Patron

JanabH.AbdulHaleem Sahib

Secretary and Correspondent

Patron

Dr.Akthar Begum

Principal

Advisor

Ms.Aariefa Basheer

Vice-Principal

Convener

Ms.R.Rashmi,H.O.D

PG Department of Computer Science and Applications

Co-Conveners

Ms.C.FirzaAfreen

Assistant Professor, PG Department of Computer Science and Applications

Ms.T.Sindhu

Assistant Professor, PG Department of Computer Science and Applications

Organizing Committee

Ms.K.Priya

Ms.T.Aruna

Ms.UmaMageshwari

Ms.T.Zeenathunnisa

Ms.F.A.Rizwana Begum

Ms.Rajalakshmi

Ms.M.Naseema

Ms.Jabeen

Ms.J.Sivasankari

Ms.I.Navitha

Ms.F.Najma

ABOUT THE COLLEGE

Islamiah Women's College was established in the year 1997. The college is permanently affiliated with Thiruvalluvar University, Vellore and approved by the Government of Tamil Nadu. It is also recognized by University Grants Commission under section 2(f) and 12(b) of the UGC Act of 1956. It is a symbol of devotion to education and love for learning. The aim of the college is to cultivate good behaviour, attitude and high values of life among its students. The students are given special training in Personality Development, Women Empowerment and in short term Value Added Courses apart from regular studies.

ABOUT THE DEPARTMENT

The PG Department of Computer Science Applications was established in 2005. The prime focus of the department is to enhance the student's ability to develop and sustain the transformative system to create future leaders. The department offers Undergraduate Courses (BCA & CS) and Post Graduate Course (CS). Excellent infrastructure, qualified and experienced faculties form the backbone of the department. The department has been successful in conducting various Workshops, Seminars and Conferences in the field of Computer Science and IT.

AIM OF THE CONFERENCE

The aim of the conference is to provide a platform to the researchers and practitioners to expose their innovative ideas and talents. The Conference provides a unique platform for all those who seek to update their knowledge.

CONTENTS

S. No	Titles	Page. No
1	APPLYING DATA MINING TECHNIQUES IN BIOINFORMATICS COMPUTING K.M.Kalaivendan & Mr.P.Magizhan	01
2	COMPREHENSIVE STUDIES AND IDEAS FOR RESEARCH IN DATA MINING TECHNIQUES P.Hariharan & Dr.K.Arulanandham	07
3	DATA MINING TECHNIQUES FOR SURVEY THE STUDENT ACADEMIC PERFORMANCE Mr.M Raghu & Mr.A.A.Khadar Maideen	11
4	DATA MINING TECHNIQUES AND APPLICATIONS USING CLUSTER ANALYSIS S.Karuna & Mr.S.Mohammed Nawaz Basha	16
5	ERADICATING OF PARTIAL DUPLICATES FROM DATA STREAMS Dhanraj Vanisree & Thomas Elizabeth Rani	20
6	IMPLEMENTATION OF MEDICAL TREATMENT PLANNER USING ADALINE T.Sindhu & R.Elango	27
7	LED LIGHT COMMUNICATION USING LIFI Ms.A.Sivasankari, Ms.S.Archana Varshini & Mrs.R.Lakshmi	33
8	SECURED DATA TRANSMISSION IN MOBILE AD-HOC NETWORKS USING MULTILEVEL ENCRYPTION TECHNIQUES P. Daniel Sundarraj & K. Arulanandam	38
9	ROLE OF GREEN TECHNOLOGY IN THE FUTURE D. Saranya & R.Shobana	43
10	CLUSTERING METHODS AND THEIR CLASSIFICATION ON IMAGES K. Priya	48
11	MANAGEMENT OF NETWORK FROM CYBER ATTACK-RANSOM WARE S.Asmathullah & V.A.Afsal Basha	53
12	NETWORK INTELLIGENCE FOR WIRELESS DATA TRANSMISSION USING SOCKET PROGRAMMING M.Dukitha, T.Balasubramanian, M.Aruna & P.Kokila	57
13	MAJOR ISSUES AND CHALLENGES ARE INVOLVE IN WIRELESS AD HOC NETWORK AND ITS SOLUTION M.Inbavalli , M.Angelinrosy, C.Anitha, S.Rekha & S.Nivetha	66

APPLYING DATA MINING TECHNIQUES IN BIOINFORMATICS COMPUTING

K.M.Kalaivendan

Research Scholar, Islamiah College (Autonomous), Thiruvalluvar University, Vellore

Mr.P.Magizhan

Research Guide, Islamiah College (Autonomous), Thiruvalluvar University, Vellore

Abstract

Bioinformatics is the application of statistics and computer science to the field of molecular biology. Its primary use since at least the late 1980s has been in genomics and genetics, particularly in those areas of genomics involving large-scale DNA sequencing. Bioinformatics now entails the creation and advancement of databases, algorithms, computational and statistical techniques and theory to solve formal and practical problems arising from the management and analysis of biological data. Over the past few decades rapid developments in genomic and other molecular research technologies and developments in information technologies have combined to produce a tremendous amount of information related to molecular biology. It is the name given to these mathematical and computing approaches used to glean understanding of biological processes. Common activities in bioinformatics include mapping and analyzing DNA and protein sequences, aligning different DNA and protein sequences to compare them and creating and viewing 3-D models of protein structures.

The primary goal of bioinformatics is to increase the understanding of biological processes. What sets it apart from other approaches, however, is its focus on developing and applying computationally intensive techniques (e.g., pattern recognition, data mining, machine learning algorithms, and visualization) to achieve this goal. Major research efforts in the field include sequence alignment, gene finding, genome assembly, drug design, drug discovery, protein structure alignment, protein structure prediction, prediction of gene expression and protein-protein interactions, genome-wide association studies and the modeling of evolution.

Introduction

Bioinformatics is an interdisciplinary research area, which may be defined as the interface between biological and computational sciences. It greatly helps in management of complex and scattered biological data, sequence analysis, algorithmic designing etc.

The term 'bioinformatics' is the short form of 'biological informatics', just as biotechnology is the short form of 'biological technology'. There are several definitions of bioinformatics, as there are for biotechnology, often depending upon 'whom, are you taking to?' Anthony Kerlavage of the Celera Genomics, defined bioinformatics as 'any application of computation to the field of biology, including data management, algorithm development, and data mining'. Clearly, a number of divergent areas, many of them outside biotechnology, come under bioinformatics.

Clearly, Bioinformatics is a multi-disciplinary research area. Different specialists involved in Bioinformatics research are Computer Scientists, Statisticians, Mathematicians, Biologists etc.

What is Bioinformatics?

Bioinformatics has emerged out of the inputs of specialists from several different areas such as biology, biochemistry, biophysics, molecular biology, biostatistics and computer science. Specially designed algorithms and organized computer databases are at the core of all bioinformatics operations. Algorithms, that are necessarily complex, make voluminous data easy to handle for defined purposes, in an amazingly short time, a process that is humanly impossible. The requirements of such an activity make heavy and high level demands on both the hardware and the software capabilities of computers.

With several divergent claimants, it is rather difficult to decide which areas of knowledge and information genuinely constitute bioinformatics. There exists a lot of confusion and sheer lack of standardization regarding what is bioinformatics and what is not. It will be pertinent to identify areas that are not normally considered as bioinformatics, as for example,

- Determination of structure of a material by crystallography and NMR,
- Ecological modeling of populations of organisms,
- Genetic mapping,
- Human structure scans,
- Artificial life simulations,
- Organism phylogenies based on non-molecular data,
- Computerized diagnosis based on pedigrees, and a few others, though all these constitute computer processing of biological data.

Bioinformatics Computing Technology

The term applies to the use of computers to store, retrieve, analyze, or predict the composition or structure of biomolecules. These include genetic materials such as nucleic acids, as well as proteins, the end of products of genes. Bioinformatics or computational biology is the use of mathematical and informational techniques, including statistics, to solve biological problems, usually by creating or using computer programs, mathematical models or both. One of the main areas of bioinformatics is the data mining and analysis of the data gathered by the various genome projects. Other areas are sequence alignment, protein structure prediction, systems biology, and virtual evolution.

The revolutionary technologies that have recently been developed to gather and analyze genomic information will help to forge a new understanding of biology, with widespread applications to medicine, agriculture, and ecology. These technologies have been made possible by developments in structural biology, engineering, and computer science, and their further advancement requires a new blend of computational analysis, micromechanical robotics, microfluidics, bioelectronics chips, imaging, and new laboratory methods for functional genomics.

Bioinformatics research is penetrating almost every sphere of human life. Some of the application areas are listed as follows:-

1. Understanding functionality of living beings so as to improve their living.
2. Drug technology
3. Identification of genetic risk factors
4. Gene therapy
5. Genetic modification of food crops and animals, etc.
6. Database administration, indexing, performance tuning
7. Software engineering
8. DNA Sequencing, Bioinformatics, Genomics, and biometrics
9. Image interpretations
10. Medical informatics
11. E-commerce and web services
12. Disaster prediction
13. Remote monitoring
14. Financial market analysis
15. Online filtering

Various Types of Computational Techniques used in Bioinformatics are as follows:

1. Knowledge discovery and machine learning
2. Algorithm design
3. Computations over graphs, constraint computation
4. Reasoning under uncertainty
5. Databases and database indexing for new data types.
6. Software and data integration.
7. Distributed computations and data use over the GRID.
8. Innovative visualization and query mechanisms for large data sets.

Systems Biology

Many scientists today refer to the next wave in bioinformatics as systems biology. Systems biology is an approach taken by scientists to tackle new and complex biological questions. Systems biology involves the integration of genomics, proteomics, and bioinformatics information to create a whole system view of a biological entity. Any system where the information can be represented digitally offers a potential application for bioinformatics. Thus bioinformatics can be applied from single cells to whole ecosystems.

The Internet vs. Bioinformatics

An important part of the field of bioinformatics is the development of new technology that enables the science of bioinformatics to proceed at a very fast pace. The internet, new software developments, new algorithms, and the development of computer cluster technology has enabled bioinformatics to make great leaps in terms of the amount of data that can be efficiently analyzed.

Data Warehouse – A New Approach to Data Collection

Data warehouse is a subject oriented, integrated, nonvolatile, time-variant collection of data in support of management's decisions. Data warehouses provide access to data for complex analysis, knowledge discovery and decision-making. They support high-level demands of organizations data and information. Several types of applications-OLAP, DSS and data mining applications are supported.

Data warehousing is defined as a process of centralized data management and retrieval. Data warehousing, like data mining, is a relatively new term although the concept itself has been around for years. Data warehousing represents an ideal vision of maintaining a central repository of all organizational data. Centralization of data is needed to maximize user access and analysis. Dramatic technological advances are making this vision a reality for many companies. And, equally dramatic advances in data analysis software are allowing users to access this data freely. The data analysis software is what supports data mining.

Data mining tools help in understanding gene expressions, drug design and other emerging problems in genomics and proteomics. Data mining also helps in performing the following tasks such as

- Gene expression analysis,
- Searching and understanding of protein mass spectroscopy data,
- 3D structural and functional analysis, mining of DNA and protein sequences for structural and functional motifs, drug design, and understanding of the origin of life
- Text mining for biological knowledge discovery.

Data ware houses have following distinctive characteristics

1. Multidimensional conceptual view.

2. Generic dimensionality
3. Unlimited dimensions and aggregation levels
4. Unrestricted cross-dimensional operations
5. Dynamic spaces
6. Client-Server architecture
7. Multi-User Support
8. Accessibility
9. Transparency
10. Intuitive data manipulations
11. Consistent reporting performance
12. Flexible reporting

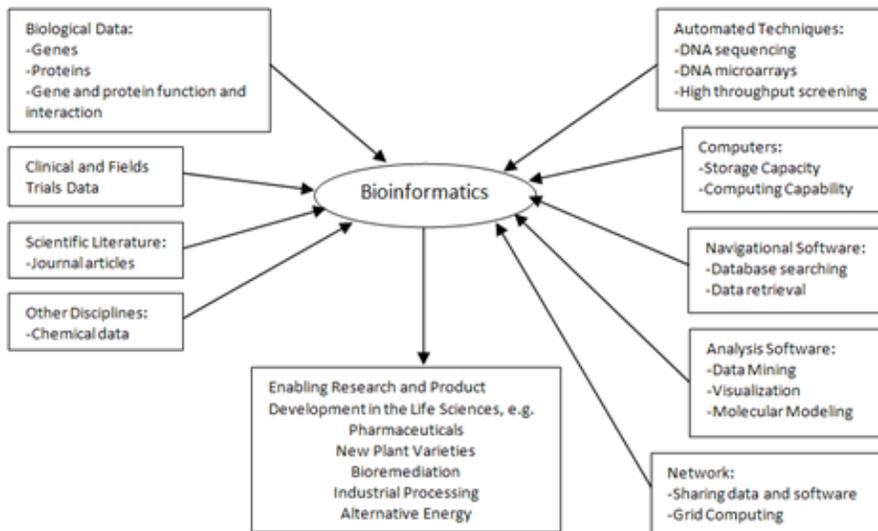


Figure: Role of IT in Bioinformatics

Bioinformatics vs. Data Mining

Before undertaking any research project in Bioinformatics the following issues should also be addressed:

1. Diversity of databases
2. Size growing exponentially and complexity also
3. Need for a common ontology and a common registration
4. Quality and provenance in databases
5. Query composition
6. Wide range of users
7. Raw data and annotation
8. Third party annotation
9. Primary and derived data
10. Query across many databases
11. How to visualize results
12. Security

Bioinformatics vs. Database

Using computers to help understand biology: Of course it is more than just putting data into a spreadsheet and making a pretty graph. We think that the real push for bioinformatics started when biologists started getting really large datasets; the kind that no longer fits well into a spreadsheet. What biologists needed were people who understood databases, and how to put data in, and more importantly get data out, so we can answer those questions that we are trying to answer. This led to the development and exponential growth of databases like GenBank. The goal of biology, in the era of the genome projects, is to develop a quantitative understanding of how living things are built from the genome that encodes them. Biological data is very complex and interlinked. A spot on a DNA array, for instance, is connected not only to immediate information about its intensity, but to layers of information about genomic location, DNA sequence, structure, function, and more. Creating information systems that allow biologists to seamlessly follow these links without getting lost in a sea of information is also a huge opportunity for computer scientists.

Database technology

Databases today can range in size into the terabytes – more than 1,000,000,000,000 bytes of data.

1. Data mining and ware housing as applied to biology
2. Data types and modeling needed for biological analysis
3. Interactive exploration and visualization for biology
4. New indexing and search structures with applications to biology

Implementation

Even with all of the public-domain databases accessible through the Internet, there will be research tasks that necessitate using a local database. The reasons vary from a need to collect, analyze, and publish sequence information inside a small laboratory to establishing a massive data warehouse as part of a pharmacogenomics R&D effort. In either case, the general issues and challenges are the same, albeit in different degrees of severity.

Infrastructure

From a hardware perspective, implementing a database requires more than servers, large hard drives, perhaps a network and the associated cables and electronics. Power conditioners and uninterruptible power supplies are needed to protect sensitive equipment and the data they contain from power surges and sudden, unplanned power outages. Providing a secure environment for data includes the usual use of username and passwords to protect accounts. However, for higher levels of assurance against data theft or manipulation, secure ID cards, dongles, and biometrics may be appropriate.

Need for Bioinformatics

Bioinformatics is the science of using information technology to further understand and enhance computational techniques in biology. The prospects are follows.

1. Assigning function to proteins
2. Mining the data
3. Comparing protein-protein interaction in different protein families
4. Transforming gene data to protein structure and correlating gene and protein function

Important advances in Biocomputing

The sequencing of the human genome has just been completed and in the next two or three years we expect progress will be made in identifying the genes. Right now, we don't even know how many genes we have! Also, in the next two or three years, we'll be learning more about the structure and function of proteins in the cell. Hopefully, in the longer term we'll be able to piece together that information to get a more complete picture of regulatory networks in the cell. Whole genome approaches. We are really good at working with the sequence of single genes at the moment. Technology is pushing towards whole genome sequencing/re-sequencing, and once we are there we will need tools to process that data. 3D protein structure prediction from amino acid sequence is another area of research where we are getting much better and could have big implications for pharmaceutical research, for example.

Conclusion

A technical set of changes faces bioinformatics and is being addressed by faster computers, technological advances in disk storage space, and increased bandwidth. One of the biggest hurdles facing bioinformatics today, however, is the small number of researchers in the field. This is changing as bioinformatics moves to the forefront of search, but this lag in expertise has led to real gaps in the knowledge of bioinformatics in the research community. Bioinformatics is about converting biological observations to a model that a computer will understand. This is very challenging task since biology can be very complex.

In short bioinformatics has revolutionized the information system in the world. Speed of flow of information, its accuracy and quality and also the management of this vast information has assisted the user to plan out their programmes most effectively to achieve their goals within set time frames. A large number of theoretical methods have been developed for sequence analysis and gene prediction. The existence of the many computational methods available for gene identification makes it very difficult to clarify the most appropriate method to be used in different kinds of analyses. The gene identification and accessing of biological database procedure is very complex due to the large database and structure of genes. Although good results have been obtained with a variety of computational approaches, the problem of gene structure prediction has not yet been completely solved. The use of these programs can significantly simplify the analysis of newly sequenced DNA, especially when applied in combination with experimental methods. The bioinformatics promises to lead to advances in understanding basic biological processes and in turn advances in the diagnosis, treatment and prevention of many genetics diseases.

References

1. Vikram Singh / Dilbag Singh / Joginder Singh.: Department of Computer Science and Technology Ch. Devi Lal University, Sirsa
2. N. Gautham.: Department of Crystallography and Biophysics University of Madras
3. Bioinformatics Computing By Bryan P.Bergeron
4. Fundamentals of Bioinformatics By S. Harisha

COMPREHENSIVE STUDIES AND IDEAS FOR RESEARCH IN DATA MINING TECHNIQUES

P.Hariharan

*Assistant Professor, PG & Research Department of Computer Science and Applications
Adhiparasakthi College of Arts & Science, G.B.Nagar, Kalavai
Gudiyatham, Vellore District, Tamilnadu, India*

Dr.K.Arulanandham

*Assistant Professor, PG & Research Department of Computer Science and Applications
Government Thirumagal Mills College, Vellore District, Tamil Nadu, India*

Abstract

Data mining is a process that uses a variety of data analysis tools to discover patterns and relationship in data that may be used to make valid predictions. Companies have used powerful computers to sift through volumes of supermarket scanner data and analyze market research reports for years however; continuous innovations in computer processing power disk storage and statistical software are dramatically increasing the accuracy of analysis while driving down the cost. Today data mining applications are available on all size systems for mainframe client/server and PC platforms. System prices range from several thousand dollars for the smallest applications up to 1 million a terabyte for the largest. Enterprise wide applications generally range in size from 10 gigabytes to over 11 terabytes. Relational database storage and management technology is adequate for many data mining applications less than 50 gigabytes. However this infrastructure needs to be significantly enhanced to support larger application. This Paper gives the idea what are the concepts and algorithms in data mining are well suited and also how to implementing the specific operation in different applications.

Introduction

Data mining only assist only business analysis with finding patterns and relationship in the data, it does not tell you the value of these patterns data mining will not automatically discover solutions without guidance rather than setting the Vague goal, "help improve the response to may direct mail solicitation " you might use data mining to find characteristics of people who (a) respond to your solicitation or (b) respond and make large purchase.

Data Mining Process

1. **Data gathering** : Data warehousing, web crawling
2. **Data cleaning**: eliminating errors and /or bogus data, example patient fever is 125 degree Celsius
3. **Feature extraction**: obtaining only the interesting attributes of the data
4. **Pattern extraction and discovery**: this is the stage that is often thought of as data mining.
5. **Evaluation of Results**: not every discovered fact is useful or even true Judgment is necessary before following your software's conclusion.

Applications Areas

Many organizations are using data mining to help manage all phases of the customer life cycle, including acquiring new customers, increasing revenue from existing customers and retaining good customers. Data mining offers values across broad spectrum of industries Telecommunications and credit card companies are two of the leaders in applying data mining to detect fraudulent use of the services. Insurance companies and stock exchange are also interest in applying this technology to reduce fraud medical applications are another fruitful area: data mining can be used to predict the effective of surgical procedures medical test or medications companies active in the financial markets

used data mining to determine market and industries characteristics as well to predict individual companies and stop performance.

3. Proposed Techniques in Data Mining for Research Work

[a] OLAP: OLAP is a part of the spectrum of decision tools traditional query and report tools describe 'what is in a database, the OLAP analyst generates a series of hypothetical patterns and relationship and uses queries against the database to verify them or disprove them OLAP analysis is essentially a detective process but what happens when a no variables of being as in the dozens or even in hundreds? It becomes much more difficult and time consuming to find a good hypothesis and analyze the database with OLAP to verify or disprove it. Data mining is different from OLAP because rather than verifying hypothetic patterns, it is essentially an inductive process take for example the analyst who wants to identify the risk factors for loan defaults using a data mining tool

[b] Apriori algorithm: The classification is implemented by using clustering algorithms and association between rules is derived by using the "apriori algorithm. We are able to find an association between flow parameters for various applications therefore making the algorithm independent of the characterized application. The rule mining helps us to various behavior patterns for an application and those behavior patterns are then fed back to refine the classification model. Traffic classification has become a crucial domain of research due to the rise in applications that are either encrypted or tend to change port consecutively. The challenge to apply the concepts of flow classification is to determine the applications involved without any information on the payload.

[c] Text mining algorithms: It is difficult for naive users to discover information pages by popular we search engine since they don't have background and domain knowledge about the status of web system. Therefore many kinds of web search engines have been developed in order to support the we information retrieval. One of the most important techniques is the text mining algorithms based on primitive association rules. Mondou provides highly relevant feedback keywords to users in order to support search steps. Using the associate keywords users can modify the combination of keywords in the initial query.

[d] Scheduling Graphs: The proposed techniques make use of novel representation named the scheduling graph which describes all possible schedules. A powerful deduction process is applied to this graph, reducing at each step the set of possible schedules. In contrast to traditional list scheduling techniques the proposed scheme tries to establish relations among instructions rather than assigning each instruction to a particular cycle. The main advantages are that wrong or poor schedules can be anticipated and discarded earlier.

[e] Text association rule mining: Text association rule mining techniques to construct term thesaurus for data preprocessing purpose. This data preprocessing is important since it has impact on the quality of other data mining techniques such as data clustering. The algorithm is applied to bibliographic databases towards the objective of enhancing traditional and content analysis. From the experiment wit a set of publications abstracts applying the proposed algorithm to combine similar terms into a pertinent concept before clustering process yields better cluster quality.

[F] SmartMobimine: The new generation of smart data mining techniques called smarmobimine that will be used in the integrated structure of the 4G mobile networks to support and develop the essential services in mobile networks. Smartmobimine techniques will integrate different types of information to support mobile users. The proposed research will be include to using Smartmobimine techniques to develop new location management, handoff, and load balancing techniques to support the integrated 4G mobile networks.

[g] Heuristic data distribution scheme: Effective data distribution techniques can significantly reduce the total execution time of a program on grid computing environments especially for data mining applications. The proposed research will be to describe a linear programming formulation for the data distribution problems on grids. HDDS will be implement parallel association rule mining method and conduct the experiments on our grid. Experimental results showed that data mining program using our HDDS to distribute could execute more efficiently than traditional scheme.

[h] k-means algorithm: Data mining techniques are broadly used in customer relationship management but there is no unified framework model for customer segmentation by now. Customer segmentation model of bank is built based on data mining which is to define the corresponding mapping relationship between customer attribute and concept attribute. The k-mean algorithm deal bank to customer segmentation and analysis the sample data which we selected from some bank. The customer shows that we can use a dynamic model of the data mining to describe customer behavior and provide useful information for the mangers of banks and decision making.

[i] Random space perturbation: In the cloud computing infrastructures, using clouds to host data query services has become an appealing solution for the advantages on scalability and cost saying. However some data might be sensitive that the data owner does not want to more to the cloud unless the data confidentiality and query privacy are guaranteed. On the other hand a secured service should still provide efficient query processing and significantly reduce the in-house workload to fully realize the benefits of cloud computing. We propose the random space perturbation (RASP) data perturbation menthod to provide secure and efficient range query and KNN query services for protected data in the cloud. The RASP data perturbation method combines order preserving encryption, dimensionality expansion, random noise injection and random projection to provide strong resilience to attacks on the perturbed data and queries.

[j] RDF: RDF is a standard model for data interchange on the web. RDF has features that facilitate data merging even of the underlying schemas differ, and it specifically supports the evolution of schemas oer time without requiring all the data consumers to be changed. RDF extends the linking structure of the web to use URIs to name the relationship between things as well as the two ends of the link. This linking structure forms a directed, labeled graph where the edges represent the named link between two resources represented by the graph nodes.

[k] Ontology concepts: E- LEARNING uses “concepts” to model the interests and preferences of a user. Since location information is important in mobile search, the concepts are further classified into two different types, namely, content concepts and location concepts. The concepts are modeled as ontology’s, in order to capture the relationships between the concepts. We observe that the characteristics of the content concepts and location concepts are different. Thus, we propose two different techniques for building the content ontology and location ontology. The ontology’s indicate a possible concept space arising from a user’s queries, which are maintained along with the Ontologies data for future preference adaptation. In E- LEARNING, we adopt ontology’s to model the concept space because they not only can represent concepts but also capture the relationships between concepts due to the different characteristics of the content concepts and location concepts.

[L] LAMSTAR: Data mining have great potential for healthcare industry to enable health systems to systematically use data and identify the efficiency and improve care with reduce cost. The data mining techniques to Multi disease treatment it can provide reliable performance. So the system can be effective in reducing the death toll. The healthcare industry collects huge amounts of healthcare data which, unfortunately are not “mined” to discover hidden information for effective decision making. Advanced data mining techniques can be helpful and can provide an efficient remedy to these kinds of

problems Large Memory Storage And Retrieval (LAMSTAR) is a method for finding the correct disease shortlisted by the doctor is obtained who confirms it by taking the necessary tests. The final report is then mined to obtain the correct symptoms. The correct symptom thus obtained is then compared with the original symptoms entered. This information is now fed to the LAMSTAR Network for assigning weights

Conclusion

The concepts of data mining are being increasingly well defined and transforming from an idea to a well defined concept with real world applications. New innovations in data mining technologies are helping increase the adoption rate across the various industries. The ontology's concepts are well suited for further to developing the relationships between concepts due to the different characteristics of the content concepts and location concepts in E - learning environment. LAMSTAR concepts have great potential for healthcare industry to enable health systems to systematically use data and identify the efficiency and improve care with reduce cost. This implementation is very much suited for further developing in health care industry.

References

1. J. Brickell and V. Shmatikov. Privacy-preserving graph algorithms in the semi-honest model. In *ASIACRYPT*, pages 236–252, 2005.
2. Shomona Gracia Jacob and R.Geetha Ramani," Discovery of Knowledge Patterns in Clinical Data through Data Mining Algorithms: Multi-class Categorization of Breast Tissue Data", International Journal of Computer Applications (0975 – 8887) Volume 32– No.7, October 2011.
3. Shweta Kharya," Using Data Mining Techniques For Diagnosis And Prognosis Of Cancer Disease", International Journal of computer Science, Engineering and Information Technology (IJCEIT), Vol.2, No.2, April 2012.
4. M.Akhil jabbar,B.L Deekshatulu and Priti Chandra," Classification of Heart Disease Using K-Nearest Neighbor and Genetic Algorithm", International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013.
5. Zhicheng Dou, Ruihua Song, JiRong Wen, "A Largescale Evaluation and Analysis of Personalized Search Strategies", 2001
6. T. Wentling, C. Waight, J. Gallaher, . La Fleur, C. Wang, &A. Kanfer, "E-learning: a review literature", Knowledge and Learning Systems Group, National Center for Supercomputing Applications, University of Illinois, 2000.
7. K. Fry, "E-learning markets and providers: some issues and respects", Education and Training, Emerland,43(4), 2001, 233-239.
8. K. Markellos, P. Markellou, M. Rigou, S. Sirmakessis, & A. Tsakalidis, web personalization experience", Proc. 5th International Conf. on Information Communication Technologies in Education, ICICTE, Samos Island, Greece, 2004.
9. A. Ben-David, N. Nisan, and B. inkas. Fairplay MP - A system for Secure multi-party computation In *CCS*, pages 257–266, 2008.

DATA MINING TECHNIQUES FOR SURVEY THE STUDENT ACADEMIC PERFORMANCE

Mr.M.Raghu

Research Scholar, Islamiah College (Autonomous), Thiruvalluvar University, Vellore

Mr.A.A.Khadar Maiden, M.C.A.

Research Guide, Islamiah College (Autonomous), Thiruvalluvar University, Vellore

Abstract

The main objective of higher education institutions is to provide quality education to its students. One way to achieve highest level of quality is to identify factors affecting academic performance and then trying to resolve weakness of these factors. The specific objective of the proposed research work is to find out if there are any patterns in the available data (student and courses records) that could be useful for predicting students' performance. The data was captured and arranged with the use of statistical package for social sciences (SPSS) and data mining tool. Developing an accurate student's performance prediction model is challenging task. This paper employs both feature reduction and classification technique to reduce error rate. In context of factors affect student academic performance, the most affecting factor is student attendance in class in addition to final exam and mid exam grades.

This document presents the architecture of the data warehouse for the architecture of three levels of a level, two levels, three classic and novel levels with their critiques and the potential areas for future work. The proposed architecture so far presents some drawbacks when applied to work in a large number of heterogeneous data sources. This paper addresses the problems of the classic data warehouse architecture and presents the typical next-generation data warehouse architecture.

Keywords: Academic performance; Prediction; Decision tree; Education data mining; Data warehouse; Scheme integration.

Introduction

Student academic performance affected by many factors like personal, social-Economic and other environmental. Knowledge about these factors and their effect on student performance can help managing their effect. Recently Educational Data Mining refers to techniques, tools, and research designed for automatically extracting meaning from large repositories of data generated by or related to people learning activities in educational environment. Predicting student's performance becomes more challenging due to the large volume of data in educational databases. The ability to predict student performance is very important in educational environments. Increasing student success is a long term goal in all academic institutions. If educational institutions can predict students' academic performance early before their final examination, then extra effort can be taken to arrange proper support for the low performing students to improve their studies and help them to success. Newly developed web-based educational technologies and the application of quality standard offer researchers' unique opportunities to study how students. The main objective of the paper is to identify both factors that affect courses success rate and student success rate then using these factors as early predictor to expected success rate and handling their weakness.

Data Mining Techniques

Data mining is an accounting method of processing data, which is successfully used in many areas to gain useful knowledge from data. Data mining techniques are used to create a model to find new knowledge information. Many large data mining techniques are created and used, including association, classification, clustering, computing, sequencing methods, and ending trees. Describing the key techniques used in the following section is as follows.

Classification

Automated classification systems automatically classify data mining based systems. This classification depends on practical applications and provides applications, programs and statistics, statistics and statistical information. Nell classification, not impairers are classified according to classification criteria.

Association

The Association is one of the best known technologies knowledge. This is witnessed, as the organization, based on the frame of the world, if he be found a customer connection is made between one quarreled over it. Therefore, there is no generation. Often combined to create the appropriate technique for two or reason and experience that the business requires. Researchers use a range of different ways of learning student's performance will come to say that such plants and irradiation of the composition and slowness of movement, Bayesian vector machine allocation, network. These include, logging timber to influence the performance of students. A decision tree is a tree in which each branch node represents a choice between a number of alternatives, and each leaf node represents a decision. Decision tree starts with a root node on which it is for users to take actions. From this node, users split each node recursively according to decision tree learning algorithm. The final result is a tree in which each branch represents a possible scenario of decision and its outcome. Among the decision tree algorithms gains popularity in terms of its higher performance in Classification accuracy.

Related Work

Many studies focus on educational data analysis to find useful information affecting quality education. Foresee future student learning behavior, sensitivity or improvement of the field, educate and educate students learning and learning and learning. Educational Data Mining (EDM) [9] is a new source of data mining, enabling the processing of minerals and equipment for data analysis in education standards. In this study Romero and Vendur's EDM detailed scientific field research, four distinct techniques, decision making mechanisms, open support systems, non-ecological neural networks, and discriminate analysis are used to build classifier models. It's used decision tree classification technique to evaluate student's performance and extract knowledge that describes students' performance in end semester examination. Their goal is to identify the dropouts and students who need special attention early before final exams.

Course Predictive Model

The work is divided into two parts: the first part is subject to a range of students' achievement. The prediction model has two main stages.

Data Collection and Preparation Phase

The data used in the study from the Social College, Computer Science and Business Administration Department at Periyar University has been collected. The total number of available data indicates the 150 Records courses from two semesters, from student decisions for 2014-2015 and 2015- 2016. 7Attributes have been selected for each subjects (Course ID, Credit Hours, and Jobs for Practical Students (if any), Number of Workplaces, Number of Exams, Final Exam Questions, Education Type, Position, Research and Success Rate). Factors based on their ability to provide acceptable computational standards are selected. Selected variables Some domain values are defined for the current trial which facilitates analytical performance. Below is a list of 1 traits and their domains in the table below. There is a unity between students and students to find out the obvious effect of success rate. In practice there

is a connection between sessions, appointments and success rates, and the results of relationships reveal. If the curriculum is part of the practice, the best from the curriculum will certainly depend on the theoretical area, the curriculum has better success rates than any other course. On the other hand, we have seen that the number of workplaces has a negative connection, as a result of student relationship; we find that most attributes are successfully associated.

Prediction Model

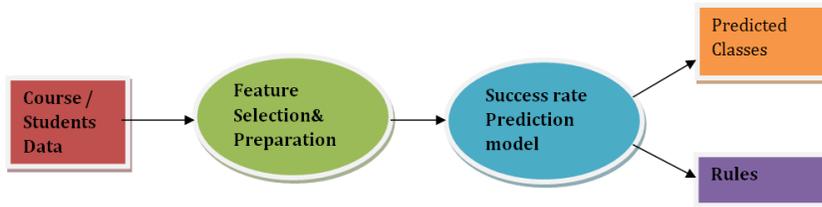


Figure 1 Prediction Model

Data Analysis Phase

Our analysis is divided into two major steps. In addition to the first step selective techniques, classification and clustering should determine the appropriate mining technique to create predictive models. Classification and benefit from previous studies findings. The algorithm classification and clustering phase selection have been selected as two step cluster techniques. The reason of selecting two-steps is its ability to divide data based on similarity without need to determining number of required clusters. This algorithm in addition to feature selection algorithm used to grantee proper attributes selection. The second step is applying the selected techniques on data then results are recorded. Both student and course data are divided into training and testing sets. Training set is labeled according to success rate. Two classes were suggested high and low. Figure 1 shows steps of course and student prediction model.

Results and Discussion

The first step in the two-step clustering algorithm and the feature selection test algorithm is certainly predicted. The facility selection technique is used to provide the most important aspects of landing in their ranks. For student data 6 featured were ranked as important and 1 as unimportant and 1 as neutral. It is important to note feature selection algorithm assign important if the rank is near or equal to 1 is this is not the case it labeled as unimportant. Supervised data are either high or low. If the success rate is greater than 65%, it is less marked. Further data is given to the classification to determine which selective traits can be used to predict the success rate. Previous Steps Student Success Rate Projection Again. Students learn the rules and student record rate. Tables 5 and 6 can be obtained from the classification process. Computational accuracy is used to calculate student average grade point average. The picture below shows some of the predictions taught from Figure 3

Data Warehouse Architecture



Figure 2 Data Warehouse Architecture

There are some improvements in the proposed approach which overcomes the drawbacks of classical three level architecture, these are:

1. In classical three- level architecture source operational databases are all grouped into one cluster and no abstraction is carried out over the associated C-Scheme.
2. The proposed structure creates an advanced data warehouse, when eliminating or modifying a savings functional database plan, the plan of clusters that comprise this database should be upgraded rather than the replacement of the fully accepted global project.
3. Analyzing the data in the classical architecture designed on a variety of source databases is very difficult. The database package of the proposed sample corresponds to each sub-program, so that is an exact semantic.
4. In this approach the metadata can be obtained during the results of the data to be reset, which will automatically remain in the compromised rank and metadata.

Typical Data Warehouse Architecture

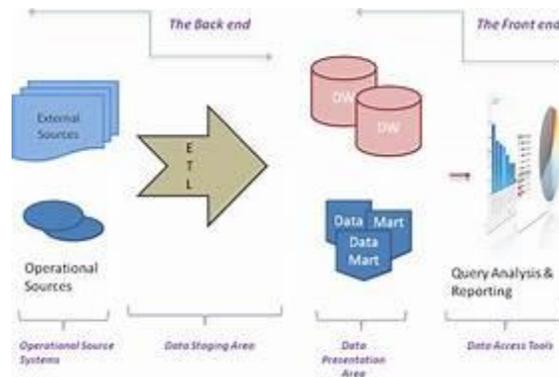


Figure 3 Typical Data Warehouse Architecture

The typical data warehouse architecture includes:

- software tools for extracting data from multiple operational databases and external data sources; for cleaning, transforming and integrating this data; for loading data into the data warehouses; and periodically refreshing the data warehouse to reflect the updates at the source.
- Here the main data warehouse consists of several data marts.
- OLAP server present multidimensional views of data to a variety of front end tools like query tool, report writer, analysis tools and data mining tools.
- Repository for storing and managing metadata.
- Tools for monitoring and administrating the warehouse system.

Critique of the Literature

- The one-level architecture induces the planning activities such as data source identification, data integration, transformation etc.
- In two-level architecture each decision support application has its own derived data; there is no possibility of storing a single copy of information for all decision support application, which indicates data duplication.
- In three-level architecture the reconciled activity eliminates inconsistencies and irregularities of operational data but again data replication found in this level. It has some, therefore they are not well related to corresponding schemes.

- The proposed novel three level architecture eliminates all the limitations presented in the classical architecture. This structure produces a semi-automatic data warehouse. This configuration is not enough to deal with structured data streams.

Conclusion

The main objective of the curriculum is to make substantial impacts on course and student performance and to show the highest efficiency of data processing applications for university management, historical data collected to deeply analyze the optimal use and techniques of data mining processes. Data mining techniques are widely used to extract previously unknown forms and to discover the connection between different aspects. On this sheet, a simple data mining system is presented on the model of the calculation. The main objective of the study is to use the classification techniques to identify students and the classification of cluster techniques in selected curriculum in order to assist academic assistants to improve their academic performance.

The proposed novel three-stage architecture in this paper eliminates the lack of classical data, which can handle structured data. Future configuration can be configured and can be configured with data to be able to extend to cope with large data. The distributed file system (such as Hadoop) is located between the source data systems and the data warehouse in such a situation. It collects, mounts, and enables large volumes of structured data, and has positions to load the data warehouse. Structured and structured data from back-end settings bring real and almost real time data warehouse.

References

Data Mining

1. W.H. Inman, DW 2.0 Architecture for the Next Generation of Data Warehousing, DM Review, Vol. 16 Issue 4, 8-25, Apr 2006.
2. W.H. Inman, What is a data warehouse?, Prism Tech Topic 1 (1), 1997.
3. K.Orr, Understanding data warehousing, Am. Program. Vol.8 (5), 2-7, 1995. www.ijircce.com

Data Warehouse

1. Baradwaj BK, Pal S (2012) Mining educational data to analyze students' performance. IJACSA 2: 63-69.
2. Nithya P, Umamaheswari B, Umadevi A (2016) A survey on educational data mining in field of education. J ComputSciSoftwDev 1: 1-6.
3. Sarker, Farhana, Thanassis T, Hugh CD (2013) Student's performance International Journal of Innovative Research in Computer. <http://www.zentut.com/data-mining/data-mining-techniques>

DATA MINING TECHNIQUES AND APPLICATIONS USING CLUSTER ANALYSIS

S.Karuna

Research Scholar, Islamiah College (Autonomous), Thiruvalluvar University, Vellore

Mr.S.Mohammed Nawaz Basha

Research Guide, Islamiah College (Autonomous), Thiruvalluvar University, Vellore

Abstract

Data mining is a process which finds useful patterns from a large amount of data. The paper discuss few of the data mining techniques, algorithms and some of the organizations which have personalized data mining technology to improve their business and found excellent outcome. Data mining is the large process has been major concern in research Community. We also spot more advantages of using several Data base management system DBMS to manage and process information converting files.

Keywords: Data mining Techniques; Data mining algorithms; Data mining applications, Classifications, Clustering, Prediction.

Overview of Data Mining

The development of Information Technology has generated a large amount of databases and enormous data in various areas. The research in database and information technology has given boost to an approach, to store and control this precious data for further decision making. Data mining is a process of pulling out of useful information and patterns from enormous data. It is also known as knowledge discovery process, knowledge mining from data, knowledge extraction or data /pattern analysis.

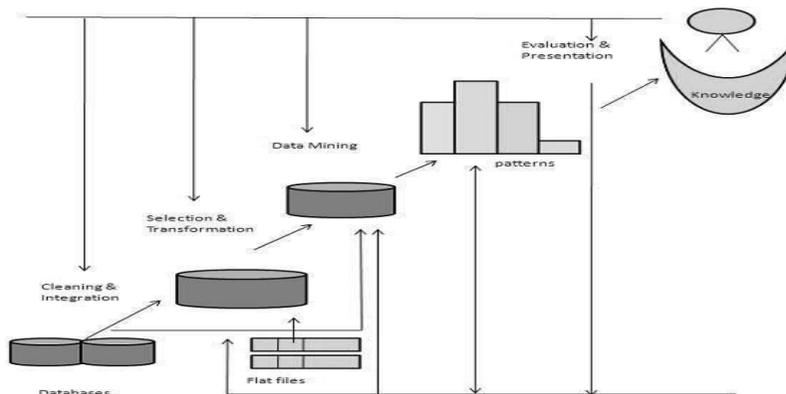


Figure: Knowledge Discovery Process

Data mining is a logical process that is used to search through a large amount of data to find the helpful data. The goal of this technique is, to find out the patterns that were previously unknown. Once these patterns are found, further they can be used to make a certain decisions for development of their business.

Three steps involved are

- 1. Exploration:** In the first step of data exploration data is cleaned and transformed into another form, and important variables and then nature of data based on the problem are determined.

2. **Pattern Identification:** Once data is explored, refined and defined for the specific variables the second step is to form pattern identification. Identify and choose the patterns which make the best prediction.
3. **Deployment:** Patterns are deployed for desired outcome.

Data Mining Algorithms and Techniques

Algorithms and techniques used for knowledge discovery from databases.

Classification

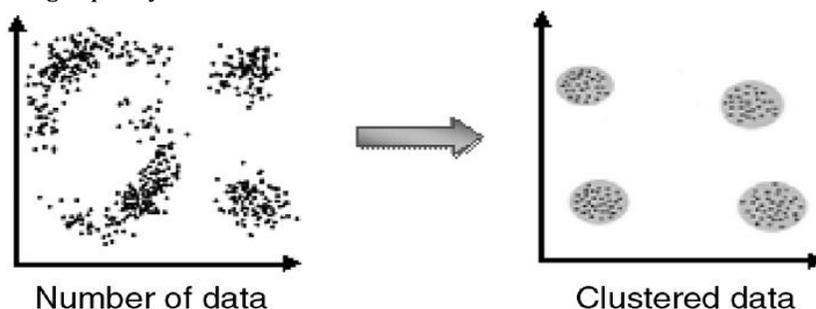
Classification is the most commonly applied data mining technique, which employs a set of pre-classified examples to develop a model that can classify the population of records at large. This leads frequently take on decision tree or neural network-based classification algorithms. The data classification process involves learning and classification. In Learning the training data are analyzed by classification algorithm. In classification test data are used to estimate the accuracy of the classification rules. If the accuracy is acceptable the rules can be applied to the new data mining. For a fraud detection application, this would include complete records of both fraudulent and valid activities determined on a record-by-record basis. The classifier-training algorithm uses these pre-arranged examples to determine the set of parameters required for proper discrimination. The algorithm then encodes these parameters into a model called a classifier.

Types of Classification Models

- Classification by decision tree induction
- Bayesian Classification
- Neural Networks
- Clustering
- Support Vector Machines (SVM)
- Classification Based on Associations

Clustering

Cluster analysis discovers clusters of data objects that are related in some logic to one another. The associates of a cluster are further like each other than they are similar to associates of other clusters. By using clustering technique we can identify intense and bare regions in entity space and can realize overall allocation pattern and correlations among data attributes. The object of clustering analysis is to find high-quality clusters.



Types of Clustering Methods

- Partitioning Methods
- Hierarchical Agglomerative (divisive) methods

- Density based methods
- Grid-based methods
- Model-based methods

Regression technique can be tailored for affirmation. Regression analysis can be used to model the relationship between one or more independent variables and dependent variables. In data mining, independent variables are attributes and retort variables are what we want to expect. Unfortunately, many real-world problems are not simply forecast. For occurrence, store prices, and product breakdown rates are all very hard to predict because they may depend on complex relations of multiple predictor variables. Therefore, more multifaceted techniques may be necessary to forecast future values. The same model types can often be used in favor of both regression and classification. For example, the CART (Classification and Regression Trees) decision tree algorithm can be used to build both classification and regression trees Neural networks too can create both classification and regression models.

Types of Regression Methods

- Linear Regression
- Multivariate Linear Regression
- Nonlinear Regression
- Multivariate Nonlinear Regression

Association Rule

Association and correlation is used to find frequent item position findings among large data positions. This type of result helps business to make certain decisions, such as catalogue design, annoyed marketing and customer shopping performance analysis. Association Rule algorithms need to generate rules with assurance values less than one. However, the number of possible Association Rules for a given data position is usually very large and a high fraction of the rules are usually of little (if any) value.

Types of Association Rule

1. Multilevel association rule
2. Multidimensional association rule
3. Quantitative association rule

Neural Networks

A neural network is a set of connected input/output units and each relationship has a weight present with it. During the knowledge phase, network learns by adjusting weights so as to be able to predict the correct class labels of the input method. Neural networks have the notable ability to obtain meaning from intricate or imprecise data and can be used to take out patterns and identify trends that are moreover complex to be noticed by either humans or other computer techniques. These are well suited for unremitting valued inputs and outputs. For example handwritten character reorganization, for training a computer to pronounce English text and many real world business problems and have already been successfully applied in many industries. Neural networks are best at identifying patterns or trends in data and well suited for prediction or forecasting needs.

Types of Neural Networks

- Back Propagation

Data Mining Applications

Data mining is a relatively new technology that has not fully matured. Despite this, there are a number of industries that are already using it on a regular basis. Some of these organizations include retail stores, hospitals, banks, and insurance companies. Many of these organizations are combining data mining with such things as statistics, pattern recognition, and other important tools. Data mining can be used to find patterns and connections that would otherwise be difficult to find. This technology is popular with many businesses because it allows them to learn more about their customers and make smart marketing decisions. Here is overview of business problems and solutions found using data mining technology.

Conclusion

Data mining has importance regarding finding the patterns, forecasting, discovery of knowledge etc., in different business domains. Data mining techniques and algorithms such as classification, clustering etc., helps in finding the patterns to decide upon the future trends in businesses to grow. Data mining has wide application domain almost in every industry where the data is generated that's why data mining is considered one of the most important frontiers in database and information systems and one of the most promising interdisciplinary developments in Information Technology.

References

1. Jiawei Han and Micheline Kamber (2006), Data Mining Concepts and Techniques, published by Morgan Kaufman, 2nd ed.
2. Dr. Gary Parker, vol 7, 2004, Data Mining: Modules in emerging fields, CD-ROM.
3. Crisp-DM 1.0 Step by step Data Mining guide from <http://www.crisp-dm.org/CRISPWP-0800.pdf>.
4. Data Mining: Techniques and Trends by N.P. Sivaselvan, B. Gopalan (Author)
5. Data Mining: Concepts and Techniques, by Jiawei Han (Author), Jian Pei (Author), Micheline Kamber (Author)

ERADICATING OF PARTIAL DUPLICATES FROM DATA STREAMS

Dhanraj Vanisree

Assistant Professor in Department of Computer Science, Don Bosco College, Yelagiri Hills

Thomas Elizabeth Rani

Assistant Professor in Department of Computer Application, Don Bosco College, Yelagiri Hills

Abstract

The data streams have a huge amount of data that be continuously inserted and queried. The oldest extant works show how to discover outliers from the rapid data. The data stream techniques allow gathering the partial data which the elements are continuously generated. The main objective of research work is to perform identifying the partial duplicate process in data streams and reduce the memory space using keywords. In this research work, Apriori algorithm is used for identifying the partial duplicate values from data streams and reduce memory space using Divide and Fold method which is a type of hashing technique, which allows a partial duplicate data insist with a keyword. It helps to reduce the memory space. The software is designed with an improved search methods of partial deduplicates with a user friendly GUI.

Introduction

Data stream is a field that is being widely studying in the area of research. Many applications have massive amount of data which store limited amount of data in storage capacity and processing time is less, thus the extraction of partial duplicates are from huge database. Traditional data stream is not suitable for extracting the partial duplicates from database so the new algorithm is regulated for this kind of extraction methods.

De Duplication

Data Deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data. Related and somewhat synonymous terms are intelligent data compression and single-instance data storage .The data Deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data. Related and somewhat synonymous terms are intelligent data compression and single-instance data storage. This technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. In the Deduplication process, unique chunks of data, or byte patterns, are identified and stored during a process of analysis. The analysis continues, other chunks are compared to the stored copy and whenever a match occurs, the redundant chunk is replaced with a small reference that points to the stored chunk. Given that the same byte pattern may occur dozens, hundreds, or even thousands of times the match frequency is dependent on the chunk size, the amount of data that must be stored or transferred can be greatly reduced.

Benefits of Deduplication

The Storage-based data Deduplication reduces the amount of storage needed for a given set of files. It is most effective in applications where many copies of very similar or even identical data are stored on a single disk a surprisingly common scenario. In the case of data backups, which routinely are performed to protect against data loss, most data in a given backup remain unchanged from the previous backup. Common backup systems try to exploit this by omitting or hard linking files that have not changed or storing differences between files. Neither approach captures all redundancies, however. Hard-linking does not help with large files that have only changed in small ways, such as an email

database. Network data Deduplication is used to reduce the number of bytes that must be transferred between endpoints, which can reduce the amount of bandwidth required. Virtual servers benefit from Deduplication because it allows nominally separate system files for each virtual server to be coalesced into a single storage space. At the same time, if a given server customizes a file, Deduplication will not change the files on the other servers something that alternatives like hard links or shared disks do not offer. Backing up or making duplicate copies of virtual environments are similarly improved.

Research Problem

Traditional data stream is not suitable for extracting the partial duplicates from database so the new algorithm is regulated for this kind of extraction methods. This research work allows partial duplicates to store in a particular field using Apriori algorithm which comes under association rule. The ASCII value has been calculated with divide and fold method and those values is mentioned using the keyword. When the keyword is titled for process then the partial duplicates will be displayed as a result. Thus it reduces memory space and time processing to produce a result; it shows a result with well-organized and well-timed. Thus it avoids unwanted data from a huge database. It avoids unwanted methods and values from database when data rapidly flows. This method is applied mostly in organizations, educational institutions, etc. There it helps to get a partial value in an efficient way and it manages memory space and processing time.

Research Objectives

The purpose of this research objective allows partial duplicates in a fixed field within a record that the data contained in database. The data that will be stored processed and accessed. This includes defining what fields of data will be stored and how that data will be stored (numeric, alphabet). This process occurs using the Apriori algorithm, which it allows the partial duplicated data from the frequent data item set. From this it scans the candidate values from the database and allows the data which are repeated in the data item set. Another method has been implied for the second process; in this stage the divide and fold method is allowed. This method involves in the selected values using some calculation based. Here, the values are divided based on the groups and providing ASCII code for each character and number, and then it had been calculated. The resulted value is considered as the place of storing the data which are partial duplicated from the item set. From this, we can able to select a required value using a keyword thus it reduce a memory space. Using of this method we can able to store a large amount of data in a data base by applying a keyword for each item set in the data base. It allows to retrieving a values in an easy way with the minimum time limit.

Research Outline

This research work allows partial duplicates to store in a particular field using Apriori algorithm which comes under association rule. The ASCII value has been calculated with divide and fold method and those values is mentioned using the keyword. When the keyword is titled for process then the partial duplicates will be displayed as a result. Thus it reduces memory space and time processing to produce a result; it shows a result with well-organized and well-timed. Thus it avoids unwanted data from a huge database.

The outcome of this research gives proficient data from database; those data are stored with valuable keyword. The keyword is formed by ASCII values; that each character is weighed by ASCII code. Thus, it reduces memory space. Each time partial values are so-called by using the keyword. This keyword helps to retrieve a partial value with appropriate results. Thus it reduces time and field

location. The memory space reduction allows other data to store in database based on ASCII code, from the data that flow from data stream.

Literature Survey and Background

Introduction

The main aim of this chapter is to describe the eradication of Deduplication software framework and relevant background of the application creation.

Different Approaches of Developing Deduplication Applications:

1. Develop the application for each language separately. Even though it is easy to host the application differently, the drawback is that keeping both the sets even after a minor change is very difficult.
2. Store all the content in different languages in database tables. This approach solves the problem to a large extent but still requires more coding and database storage. This can have a performance issue also as the content is fetched from the database; but has the advantage of easy maintainability and also creation of the language.
3. Store all the content in different languages in resource files. Resource is a non-executable data that is required by the application and is deployed along with the application. This will have a better performance compared to fetching from the database; but deployment and new language support can be little cumbersome as it requires more technical knowledge.

Database

A database is a collection of information that is organized so that it can easily be accessed, managed and updated. But it is quite difficult to store and manipulate more than one language in database.

MSSQL

Microsoft SQL Server is a relational database management system developed by Microsoft. As a database, it is a software product whose primary function is to store and retrieve data as requested by other software applications on the same computer or those running on another computer across a network (including the Internet). There are at least a dozen different editions of Microsoft SQL Server aimed at different audiences and for workloads ranging from small single-machine applications to large Internet-facing applications with many concurrent users.

Microsoft Visual Studio

Microsoft Visual Studio is an integrated development environment (IDE) from Microsoft. It is used to develop computer programs for Microsoft Windows super family of operating systems, as well as web sites, web applications and web services. Visual Studio uses Microsoft software development platforms such as Windows API, Windows Forms, and Microsoft Silver light. It can produce both native code and managed code.

Windows Forms Designer

The Windows Forms designer is used to build GUI applications using Windows Forms. Layout can be controlled by housing the controls inside other containers or locking them to the side of the form. Controls that display data (like textbox, list box, grid view, etc.) can be bound to data sources like databases or queries. Data-bound controls can be created by dragging items from the Data Sources window onto a design surface. The UI is linked with code using an event-driven programming model. The designer generates either C# or ASP.NET code for the application.

Information of Deduplication

Data stream is a field that is being widely studying in the area of research. Many applications have massive amount of data which store limited amount of data in storage capacity and processing time is less, thus the extraction of partial duplicates are from huge database. Traditional data stream is not suitable for extracting the partial duplicates from database so the new algorithm is regulated for this kind of extraction methods.

This research work allows partial duplicates to store in a particular field using Apriori algorithm which comes under association rule. The ASCII value has been calculated with divide and fold method and those values is mentioned using the keyword. When the keyword is titled for process then the partial duplicates will be displayed as a result. Thus it reduces memory space and time processing to produce a result; it shows a result with well-organized and well-timed. Thus it avoids unwanted data from a huge database.

This method is applied for allowing partial duplicates from huge database with a new algorithm. The data stream is one of the methods to finding outliers. This research objective focuses only on gathering partial duplicates from data. It avoids unwanted methods and values from database when data rapidly flows. This method is applied mostly in organizations, educational institutions, etc. There it helps to get a partial value in an efficient way and it manages memory space and processing time.

The outcome of this research gives proficient data from database; those data are stored with valuable keyword. The keyword is formed by ASCII values; that each character is weighed by

ASCII code. Thus, it reduces memory space. Each time partial values are so-called by using the keyword. This keyword helps to retrieve a partial value with appropriate results. Thus it reduces time and field location. The memory space reduction allows other data to store in database based on ASCII code, from the data that flow from data stream.

Method Advantages

1. The purpose of this research objective allows partial duplicates in a fixed field within a record that the data contained in database. The data that will be stored processed and accessed. This includes defining what fields of data will be stored and how that data will be stored (numeric, alphabet). This process occurs using the Apriori algorithm, which it allows the partial duplicated data from the frequent data item set. From this it scans the candidate values from the database and allows the data which are repeated in the data item set. Another method has been implied for the second process; in this stage the divide and fold method is allowed. This method involves in the selected values using some calculation based. Here, the values are divided based on the groups and providing ASCII code for each character and number, and then it had been calculated. The resulted value is considered as the place of storing the data which are partial duplicated from the item set. From this, we can able to select a required value using a keyword thus it reduce a memory space. Using of this method we can able to store a large amount of data in a data base by applying a keyword for each item set in the data base. It allows to retrieving a values in an easy way with the minimum time limit.
2. For many organizations it's a big challenge to archive or retrieve data while providing users with access to their information when they need it. By relocating inactive data from expensive production system and legacy database while preserving data integrity and access. The complex problem arises when an organization has multiple sources of data. To avoid this complexity in an organization this method is applied to maintain the huge database and it allow the users to retrieve data when they need it. Then storing of data will make easiest way. In this research process, allow

- the users to store the data by using a unique keyword, from this keyword the users can easily retrieve whenever they need it.
3. This method is applied for allowing partial duplicates from huge database with a new algorithm. It applied mostly for organizations, educational institutions, etc., there it could contained large amount of data. Thus it helps the user to retrieve a particular data from the database using a keyword. This reduces memory space and processing time to store such other data according to the keyword.
 4. The outcomes of this research process which provide capable results from database. It's work because of the Apriori algorithm and divide and fold. The algorithm allows calculating the partial duplicated values from data base, and stored in a separate block which are repeating in process. Then the collected values are calculated using the divide and fold method where the ASCII code are applied for each value, the final resultant field is considered to store the partial duplicate item set in the data base and it can be retrieved and stored using a unique keyword.

Method Description

The Eradicating of partial duplicates from data stream is a method which used to remove the partial duplicates from a given data set and stored in a particular area. This method helps to remove the partial duplicates from the given data set and retrieve those data item using the keyword. The eradication of partial duplicates from data set is using the Apriori Algorithm which is depending on the association rule. The data item are denoted by keyword using hashing technique i.e., divide and fold method. The methods are as follows: The method involves in this research area is based on the two algorithms. These algorithms allow gathering partial DE duplicate values from database and store those values by assigning a keyword.

Phase1_ Apriori Algorithm

Apriori algorithm can be considered as a primitive method in this area. It is suitable for large database so it has been applied for data streaming. This method introduces a concept of assigning a keyword. Based on the keyword it overcomes the difficulties from data stream. It works based on the following steps: First it scans candidate item set of size (K), then it scans the frequent item set of size (Lk) in the database. After classifying those items according to the support values, it removes the least values from the collection. The remaining values are listed in the table. Now the items are calculated based on (Ck+1) for each transaction 't' in database increment the count of all candidates in Ck+1 contained in 't' from the remaining item set. Again it scans for items then it removes least value. This process is continued until the partial values are grouped into single sets.

Ck: Candidate item set of size K

Lk: Frequent item set of size K

L1: {Frequent item set} ;

for (k=1;Lk!=E;K++)do Ck+1=Candidates generated from Lk; for each transaction 't' in db do
increment the count of all candidates in Ck+1 contained in t;

end for

Lk+1= Candidates in Ck+1 min_support

End for

Return UkLk

Phase2_Divide and Fold Method

This is another method in this research area. Here the ASCII code is used for each character and number in this process. A calculation is made using the ASCII code to find out the field where the partial duplicates values are stored.

The Example as Follows

Ex 1:

Using Apriori Algorithm

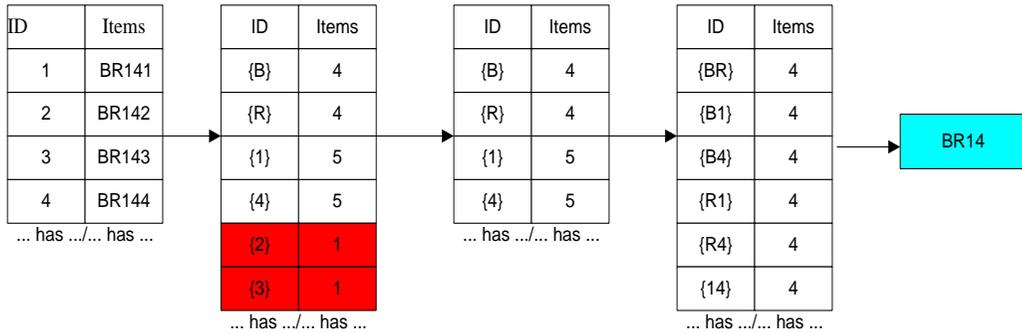


Table 2.1 Ex.1 Apriori Algorithm

The step involves in Apriori algorithm as follows:

1. First it allows the data to occupy the database rapidly [BR141, BR142, BR143, BR144].
2. It scans the size of candidate values from item set [6].
3. It checks the size of an each frequent item set [B (4), R (4), 1(5), 4 (5), 2 (1), 3 (1)].
4. If the frequent item set contained least value then it removes from the set [2 (1), 3 (1)].
5. Then again scans the size of candidate values from item set; checks the values which are frequently available. If the item

Results and Discussion

Introduction

This chapter states the results obtained by the proposed framework. The deduplication is viewed by the GUI that is implemented in ASP.Net Web application. The deduplication software framework is useful for the user who would like to remove the partial values. The proposed framework retrieves the results from the MySQL database based on the user’s query and displays the results. The software is tested with the end user and the results are evaluated.

Future Enhancements

The Deduplication database software framework exhibits the development of information retrieval without any user constraints. In future the framework can be reused with some modification for some other categories. The same application can also be reused by adding the other information’s.

Conclusion

A software framework for a deduplication database application is presented in detail. The framework is evaluated by developing a desktop application for Eradication of partial duplicates using ASP.Net Web Application as front end and MSSQL database as backend. The application is portable and the user can install and access the application in any system. The GUI designed in the application is

simple and user friendly. Effective search methods are incorporated in the application to search the data from database based on a criteria.

In this research work, a simple demonstration is shown for data streaming. When the data flows rapidly from database it retrieves the partial duplicate values. A new method and algorithm are proposed for this research to gather the partial duplicate values and remove the unwanted data from database. The limitations have been fixed when selecting the values from database. Thus, the results will be produced efficiently and there is no wastage in database when the values are entered. Each value will be fixed to particular field based on the ASCII code(Apriori Algorithm and Hashing Technique) used in this process

References

1. Dr. S. Vijayarani¹, Ms. P. Jothi "An Efficient Clustering Algorithm for Outliers Detection in Data Streams" International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 9, September 2013
2. ShirinMohammadi, Ali A. Safaei, FatemehAbdi and Mostafa S. Haghjoo," Adaptive Data Stream Management System Using Learning Automata" Advanced Computing: An International Journal (ACIJ), Vol.2, No.5, September 2011
3. MadjidKhalilian,Norwati Mustapha, "Data Stream Clustering: Challenges and Issues" Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol I IMECS 2010 march 17-19 2010 Hong Kong
4. DikshaUpadhyay, Susheel Jain, Anurag Jain, "Comparative Analysis of Various Data Stream Mining Procedure and Various Dimension Reduction Techniques", ISSN 2224-5774 (Paper) ISSN 2225-0492 (Online) Vol.3, No.4, 2013
5. YixinChen,LiTu," Density-Based Clustering for Real-Time Stream Data"
6. HosseinMoradiKoupaie, Suhaimi Ibrahim, JavadHosseinKhani," Outlier Detection in Stream Data by Machine Learning and Feature Selection Methods", International Journal of Advanced Computer Science and Information Technology (IJACSIT)Vol. 2, No. 3, 2013, Page: 17-24, ISSN: 2296-1739
7. PrakashChandore, PrashantChatur," Outlier DetectionTechniques over Streaming Data in Data Mining: A Research Perspective", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-1, March 2013
8. H.S.Behera,AbhishekGhosh,Sipakku. Mishra,"A New Hybridized K-Means Clustering Based Outlier Detection Technique For Effective Data Mining" Volume 2, Issue 4, April 2012 ISSN: 2277 128X
9. Neeraj And Anuradha, "Mining Maximum Frequent Item Sets Over Data Streams Using Transaction Sliding Window Techniques" International Journal of Information Technology Convergence and Services (IJITCS) Vol.3, No.2, April 2013
10. MHMOOD DEYPIR, MOHAMMAD HADI SADREDDINI AND MEHRAN TARAHOMI," An Efficient Sliding Window Based Algorithm for Adaptive Frequent Itemset Mining over Data Streams" JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 29, 1001-1020 (2013).
11. DikshaUpadhyay, Susheel Jain, Anurag Jain "Comparative Analysis of Various Data Stream Mining Procedures and Various Dimension Reduction Techniques" ISSN 2224-5774 (Paper) ISSN 2225-0492 (Online) Vol.3, No.4, 2013
12. <http://www.ats.ucla.edu/stat/sas/library/nesug99/ss123.pdf>
13. <http://www.mimuw.edu.pl/~son/datamining/DM/4-preprocess.pdf>

IMPLEMENTATION OF MEDICAL TREATMENT PLANNER USING ADALINE

Ms.T.Sindhu

Assistant Professor, Islamiah Women's College, Vaniyambadi

Mr.R.Elango

Guest Lecturer, Government Arts College for Men, Krishnagiri

Abstract

The aim of the proposed research work is to estimate the length of stay of patients in a hospital by considering the symptoms of their diseases. The length of stay has a direct relationship with the severity of illness. In this proposed paper ADALINE neural network is implemented using MATLAB over the data (represents the symptoms of the patients) for analysis and it is categorized, based on classification done by supervised learning techniques. The eight attributes used (representing the symptoms of the patients) are given as inputs under the ADALINE neural network. The major objective of this research work is to achieve an automated categorization, Which specifies the length of stay of patients depending on the severity of illness. The proposed techniques is cost effective in doing the categorization when compare to the existing manual method.

Introduction

Learning Algorithm

The assumption made in [1] is,

η is the learning rate (some constant)

y is the desired output

\hat{y} is the actual output

Then the weights are updated as follows $w \leftarrow w + \eta(y - \hat{y})x$. The ADALINE converges to the least squares error which is $E = (\hat{y} - y)^2$. This update rule is in fact the stochastic gradient descent update for linear regression.

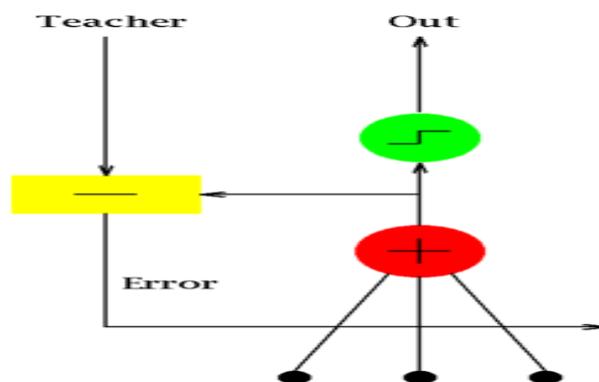


Figure Learning Inside a Single Layer ADALINE

An effective ADALINE Network is a training of multi-layer perceptrons depends on the Incorporation of an appropriate error or objective function. Classification-based (CB) error functions are heuristic approaches that attempt to guide the network directly to correct pattern Classification

rather than using common error minimization heuristics, such as sum-squared error and cross-entropy, which do not explicitly minimize classification error. The artificial neural network (ANN) was inspired by biological models of neurological systems and is an established machine learning model with robust learning properties and simple deployment. ANNs are often used as a “black box” that receives data observations as input and outputs decisions based on these observations. This work focuses on studying and improving the theoretical and practical use of artificial neural networks in classification problem domains

ADALINE Network

ADALINE (Adaptive Linear Neuron or later **Adaptive Linear Element**) is an early single-layer neural network and the name of the physical device that implemented this network. It was developed by Professor Bernard Widrow and his graduate student Ted Hoff at Stanford University in 1960. It is based on the according to, as in McCulloch–Pitts neuron. It consists of a weight, a bias and a summation function. As in [1] ADALINE network multiple epochs with data from a sliding window of the system’s input output data. Simulation results show that the proposed method provides a much faster convergence speed and better tracking of time varying parameters. The low computational complexity makes this method suitable for online system identification and real time adaptive control

applications. As in[2] neural network for handwriting recognition, a set of input neurons may be activated by the pixels of an input image representing a letter or digit. The activations of these neurons are then passed on, weighted and transformed by some function determined by the network's designer, to other neurons, etc., until finally an output neuron is activated that determines which character was read. As in [3] possibility of predicting average summer-monsoon rainfall over India has been analyzed through Artificial Neural Network models. In formulating the Artificial Neural Network based predictive model, three layered networks have been constructed with sigmoid non-linearity. Node accepts multiple inputs and generates one output. Given the following variables:

- x is the input vector
- w is the weight vector
- n is the number of inputs
- θ some constant
- y is the output

$$y = \sum_{j=1}^n x_j w_j + \theta$$

- then we find that the output is

If we further assume that

- $x_{n+1} = 1$
- $w_{n+1} = \theta$

then the o/p reduces to the dot product of x and w $y = x \cdot w$

Learning Techniques in Neural Networks

There are three major learning paradigms, each corresponding to a particular abstract learning task. These are

- Supervised learning,
- Unsupervised learning
- Reinforcement learning.
- Decision tree learning.

Proposed

The ADALINE network works quite well for many Applications, but it is restricted to a linear problem Space. The input patterns in the ADALINE’S training set must be linearly separable otherwise

the ADALINE will never categorize all of the training patterns that for all below the threshold value. however the ADALINE is guaranteed to reach its minimum error since there are no obstacles along the error surface to interface with the training process. Training occurs by repeatedly presenting sets of data composed of input pattern, where the ADALINE minimizes the number of errors it makes when sorting the patterns into their correct categories. Once trained, the ADALINE can categorize new input attributes according to the experience it gained.

Problem Description

The data contains 500 patterns observing the symptoms of the patient admitted to the hospital. Each pattern is a set of 8 attributes (Complications, Diseases of blood, Body systems involved, Diabetes, Hyper tension, Obesity, Smoker, Anatomical surgical groups) in the range 0.0 to 1.0 . Each attribute represents a particular symptom faced by a patient suffering from a disease. The three classes are represented as 0,1 and 2, where 0 denotes minimum stay of the patient in the hospital, 1 denotes the medium number of day to the hospital and 2 denotes the maximum number of days of stay.

Assumption

ADALINE IS USED TO PREDICT THE SEVERITY OF ILLNESS of accordingly propose the use of hospital resources. Expenses will be reduced, by avoiding unnecessary tests and procedures, lowering the length of stays and procedural changes. Individually trained neural network learn how to classify and predict the severity of illness for particular diagnoses, so that the quality and cost issues can be addressed fairly. In this thesis lowering the length of stay, depending upon the severity of illness is focused.

The input variables addressing the eight attributes associated with a patient, which are considered in this thesis are:

1. Complications
2. Diseases of blood
3. Body systems involved
4. Diabetes
5. Hyper tension
6. Obesity
7. Smoker
8. Anatomical surgical groups

The output variables that help the patient is knowing their length of stay, is subdivided in to three categories.

1. Minimum-1 to5 Days for class label is 0
2. Medium-5 to15 Days for class label is 1
3. Maximum-15 to30 Days for class label is 2

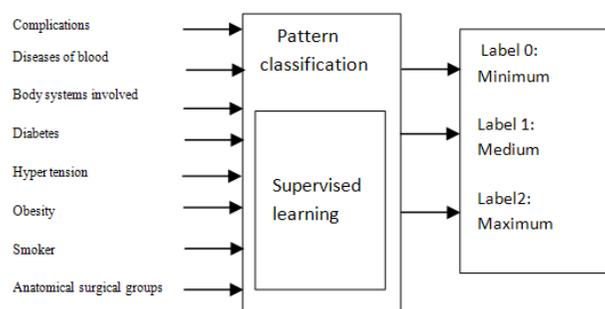


Figure Block Diagram

Results and Discussion

Dataset

	0	1	2	3	4	5	6	7	8	9
0	0.35	0.265	0.09	0.2255	0.0995	0.0485	0.07	0.77		
1	0.05	0.57	0.63	0.1735	0.4975	0.6405	0.845	0.61		
10	0.043	0.42	0.135	0.277	0.2565	0.1415	0.51	0.29		
11	0.6	0.96	0.65	0.615	0.85	0.775	0.34	0.65		
12	0.32	0.515	0.514	0.24	0.3105	0.675	0.705	0.1		
20	0.56	0.85	0.75	0.71	0.835	0.6	0.87	0.15		
21	0.44	0.365	0.125	0.416	0.2155	0.214	0.55	0.61		
22	0.96	0.35	0.12	0.515	0.224	0.108	0.1565	0.9		
30	0.56	0.95	0.85	0.67	0.805	0.75	0.641	0.9		
31	0.88	0.205	0.08	0.77	0.052	0.039	0.62	0.3		
33	0.33	0.255	0.08	0.205	0.195	0.95	0.35	0.37		
40	0.353	0.415	0.15	0.3775	0.237	0.615	0.533	0.2		
41	0.025	0.95	0.165	0.62	0.507	0.618	0.39	0.1		
110	0.165	0.535	0.195	0.606	0.5755	0.88	0.58	0.14		
111	0.425	0.3	0.095	0.3515	0.141	0.775	0.12	0.8		
120	0.615	0.21	0.86	0.225	0.76	0.75	0.635	0.5		
121	0.515	0.245	0.85	0.235	0.53	0.0475	0.05	0.48		
130	0.245	0.425	0.125	0.368	0.294	0.95	0.626	0.6		
140	0.49	0.38	0.335	0.415	0.9175	0.095	0.9	0.71		

Figure 1 Data Set

The class member will denotes in the first column to identify the class. Each patient has an eight attribute values, it will take from the symptoms of the patients.

Classification Based On Threshold Value

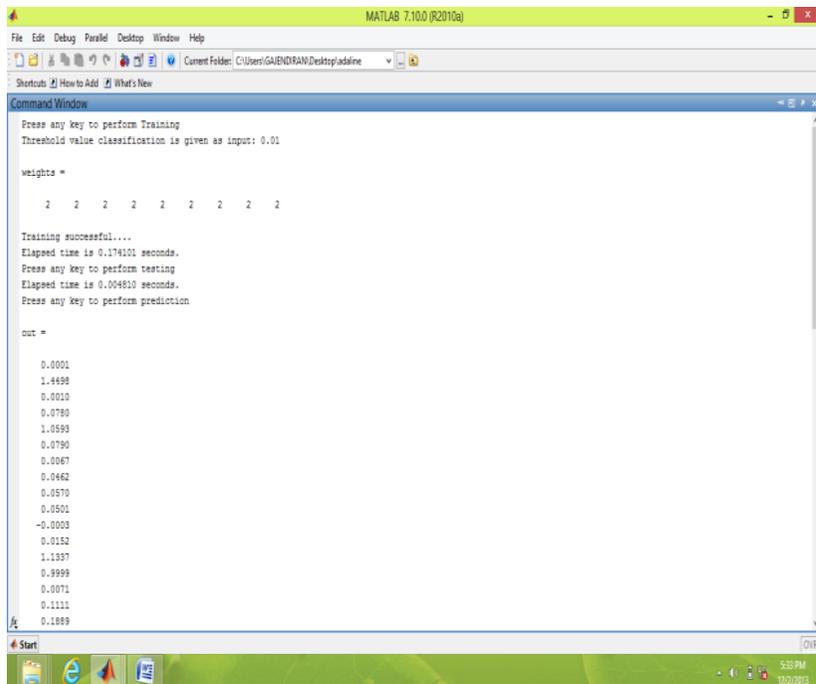
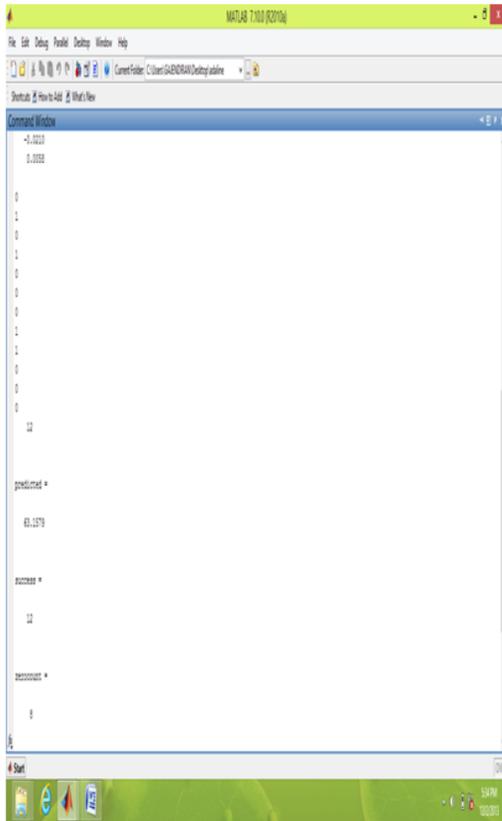


Figure 2 Classification based on Threshold Value

Enter the threshold value that value only fixing the range. In which patient's details will match to the threshold value that value will filter separately for classification process.

Classification of Predicted Values



Status of Staying Patients

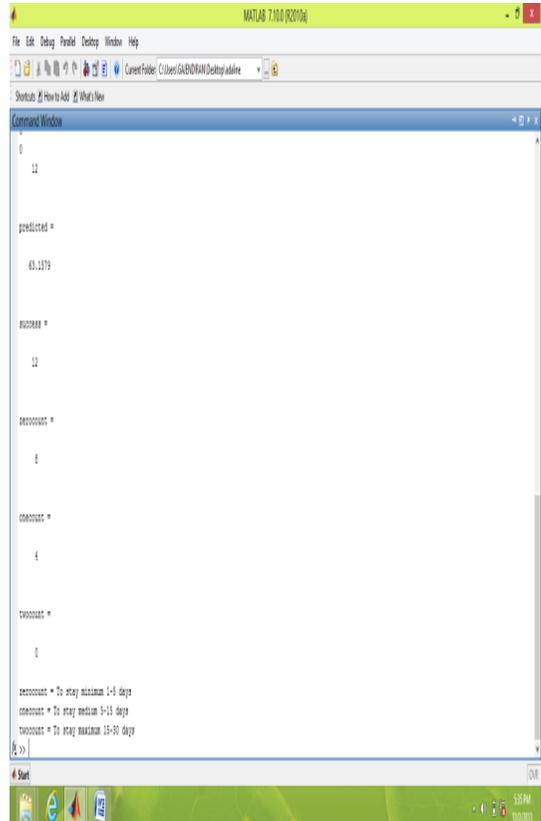


Figure 3 Classification of Predicted value, Figure 4 Status of the staying patient

The predicted values only classified into the three class members. The class members are zero count, one count and two counts. These counts are coming under the pattern classification. The patient under which categories (Member) and then the prediction is how long they will be stay in the hospital.

Classification of Actual Class Member Predict Count Vs Actual Count

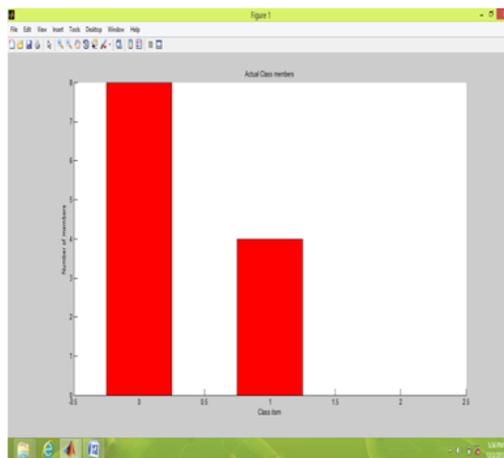


Figure 5 Graph Representation of Actual Class Member

This figure will display the diagrammatical representation of the class member in which categories the patient is coming, that representation will marked into red colour.

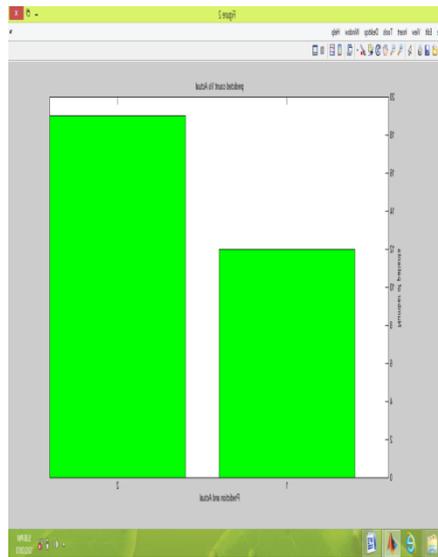


Figure 6 Graph Representation of Predict and Actual Count Value

This figure will compare the predicted and actual values. It will show the difference between predict and actual values.

Conclusion

This research work have proposed ADALINE neural network was implemented for pattern classification with a collection of information containing the data of symptoms of the patient admitted in the hospital. The main aim of this research work is to assist and estimate the length of staying at in a hospital by considering the symptoms of their diseases. The length of stay has direct relationship of the severity of illness. The classification has three categories such as minimum, medium, maximum. An each pattern is a set of eight attributes as a input in the range of 0.0 to 1.0 and which represents the particular symptoms faced by a patient who has suffering from the diseases. The three classes are represented as 0,1and 2. The class 0 denotes minimum stay (1-5 days) , class 1 denotes medium stay (5-15) and class 2 denotes maximum stay (15-30 days) of the patients stay in the hospital. So this research work for “implementation of medical treatment planner using adaline” is very much helpful to hospital and it will reduce the system time for developed system.

References

1. Bernad Widrow, (1980). “**ADALINE (Adaptive Linear Neuron or later Adaptive Linear Element) is an early single-layer neural network**”
2. Dr. Zhang, (2001). “system identification and control, neural networks and intelligent control, discrete event dynamic systems and manufacturing systems”.
3. Frank Rosenblatt, (1958). “created the **perceptron**, an algorithm for **pattern recognition** based on a **Neural network**”
4. S Chattopadhyay, ((2006. “possibility of predicting average summer-monsoon rainfall over India has been analyzed through Artificial Neural Network models”
5. A Konar ,2005. “the scope of application of Widrow-Hoff’s ADALINE model from binary to gray level (fuzzy) pattern recognition
6. Hinton, G. E.; Osindero, S.; Teh, Y. (2006). "A fast learning algorithm for deep belief nets".*Neural Computation*.
7. Fukushima, K. (1980). "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position". *Biological Cybernetics*. 8. M Riesenhuber, T Poggio. Hierarchical models of object recognition in cortex. *Nature neuroscience*, 1999

LED LIGHT COMMUNICATION USING LIFI

Ms.A.Sivasankari

*Head of the Department, Department of Computer Science and Applications
D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India*

Ms.S.Archana Varshini

*Research Scholar, Department of Computer Science and Applications
D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India*

Mrs.R.Lakshmi

*Assistant Professor, Department of Computer Science and Applications
D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India*

Abstract

In this generation of advanced science, a number of wireless connections are available to connect to the internet. It has become a very common and a useful tool to access wireless internet. As the number of people accessing the wireless internet increase, the blocked airwaves make it extremely difficult to latch on to a reliable signal. Wi-Fi may not be the most efficient way to provide new desired capabilities such as precision indoor position and signal recognition. Li-Fi is mostly recently referred to as optical wireless technology may also called as visible light communication (VLC) on the other hand, suggest an absolutely new prototype in wireless technologies in terms of communication speed, flexibility and usability.

Li-Fi is some have used to tag the fast and cheap wireless-communication system which is the optical version of WiFi. LED lights flicker (constant fluctuation of light output from on to off) at extremely high speed that human eyes can't detect. Binary information is made up of strings of 1's and 0's. Any light source can transmit this ON and OFF information but LEDs are capable of highest glittering speed. The Li-Fi work when the device sense the light that being emitted by the Li-Fi router, meaning it will only work if you're in the same room or area the light is being emitted and do not require Line-Of-Sight(the straight line between source and target) conditions. More sophisticated techniques could dramatically increase VLC data rate.

Keyword: *LI-FI (light fidelity), VLC (visible light communication), LED (Light Emitting Diode), WI-FI (wireless fidelity)*

Introduction

The wireless communication and the network which face the difficulty of access speed, when the various devices is connected. It creates a common difficulty when user connected to the multiple users, the permanent bandwidth which is existing for the particular user makes it more difficult to reveal high data transfer rates and connect to a secure network. The term Li-Fi was coined by Harald Haas and it's a form of light communication and a subset of optical light communication (OLC) and could be a balance to Radio Frequency (RF) communication(WI-FI or Cellular network), or even a replacement in contexts of data broadcasting. The measurement of Li-Fi is about 100 times faster than some Wi-Fi implementations, reaching speeds of 224 gigabits per second. The Li-Fi is wireless and uses infra-red and or visible light communication which is near to ultraviolet spectrum, that is part of optical wireless communications knowledge, which carries a lot information, and has been planned as a solution to the Radio Frequency bandwidth restrictions . The physicist Harald Hass has come up with a solution which he names as "data through illumination" taking the fiber out of fiber optics by sending data through LED light bulb that vary in intensity faster than a human eye and can identify. This technology is named as Li-Fi transmission which stands for light fidelity. This invention can produce data rates faster than 10mega bits per second which is much more than that of an average broadband connection speed. The use of light in order to broadcast data can be insufficient and association to radio waves, there is a great amount of possibilities that can be developed due to this technology. In

fundamental nature of a single pixel that could transmit a single channel of information to a source. The importance of Li-Fi technology are distance, cost, traffic updates, television interaction and game console. The technology is still in its infant stages, the usefulness of this Li-Fi technology has implications for a great amount of quality.

Optical Wireless Field Instruments

Optical wireless transmitters are constituted by a sensor for real-time measurement of process variable information and the timely transmission per update rate of OWI to the OAP via optical signal. Optical wireless control elements are constituted by the control element and the light-sensitive module, which receive the optical signal at the light-sensitive sensor and convert it into a respective electrical signal to control the process variable. Optical access point (OAPs)/transceiver modules comprise LED lights, light-sensitive sensors such as photo-detector or solar-panel, high-speed LED drivers, amplification and processing modules and power modules. The light-sensitive sensor is connected to an optical receiver or signal processing and amplification module, which processes the process-variable information transmitted from the optical wireless transmitters and sends to the DCS/PLC in the LCR/LER/SUB via redundant fiberoptic cable. The switch control module converts the control signal from the DCS/PLC into a switching signal of a specific frequency. The high-speed LED driver receives the switch signal, which is converted to LED flashing lights—optical signal. Intrinsically safe battery packs or renewable energy, such as thermal, solar, vibration/wind, power the light-sensitivity sensor module, LED array and processing module.

Issues in Wifi Communication

(1) Speed:

Through 802.11ac tentatively supports an Ethernet-corresponding speed of one gigabit per second, or 1Gbps, as of 2013 feb, the a large amount accepted 802.11n standard wires a greatest speed of 600Mbps a bit more than half the speed of 802.11ac. In real-world applications, wireless networking is significantly slower than wired networking when transferring files over a local area network. The problem becomes even more sharp if user have a hard network or if user signal power is weaker than optimal.

(2) Reliability

Wireless networking uses radios to transmit signals. The global or satellite radio, wireless networks have a partial number of channels and, if every channel is filled, connections will slow down or break down to work. The radios that make Wi-Fi work are also horizontal to interference. Cell phones, microwave ovens, walls, and large pieces of metal like those that make up filing cabinets can all interfere with Wi-Fi signals, giving user unreliable network performance.

(3) Security

A wired network connection can only be intercept by somebody who has spliced into the wire. But because wireless connections go through the air, all that a person interested in stealing your information needs is a wi-fi receivers, software, patience, and a place to work where he can receive your signal. While the wi-fi protected access security protocol is better than nothing, it can still be cracked by a dedicated hacker.

(4) Reduced management control

Opening up user organizational network that reduces user control in two ways. The first is that it makes it easier for user employees to connect their own non certified devices, like personal tablets

or cell phones, to your network, consuming your bandwidth. The second is that it allows them to work wherever they want, as long as they are within range of the signal.

Implementation of Li-Fi technology

The implementation of light-fidelity (Li-Fi) eliminates the use of instrument wiring systems and cable infrastructure, thereby reducing the cost and security risks. The Li-Fi instrument communication networks are best-suited for plants where Wi-Fi and other radiation is bad for sensitive areas. It offers safe, abundant connectivity and can solve issues related to the short range of radio-frequency bandwidth. The Li-Fi instrument communication network also could enable greater capabilities to realize the Industrial Internet of Things (IIoT) and offer advanced connectivity between field devices and systems that go beyond machine-to-machine connections. The system has relatively low capital expenditure (CAPEX) and operating expenses (OPEX).

Building the Modular Li-Fi Instrument Communication Network Requires these Steps

Step 1: The transmitter feeler module like temperature, pressure, level, flow measures the real-time process variables in the plant. The real-time process-variable information will be sent to the nearby optical access point (OAP) via optical signal that transmitted light, not the radio signal. The process variable information shall be updated to the OAP or DCS, as per the user-defined update rate. The update rate of the optical wireless instruments will be decided on the basis of loop criticality of the process.

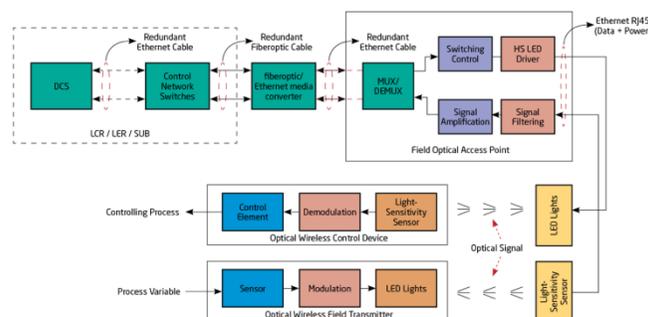
The optical transmitter module measures the real-time process-variable information, and the process-variable value is sent to the transmitter control module via hardwire signal.

The transmitter control module processes the process-variable value and converts it into a certain frequency of the switching signal. This encoded switching signal drives the high-speed LED driver, which controls the switching of LED lights to create the optical signal transmitted to the corresponding optical access points (OAPs).

Step 2: The light-sensitivity sensor of the OAP, senses the flickering or change in intensity modulation or on-off frequency of LED lights and converts the optical signal from the field optical transmitter into a corresponding electrical signal. Then the output of the light-sensitive sensor is sent to a signal processing and amplification module.

Step 3: The signal-filtering, -processing and -amplification module receive the output electrical signal from the light-sensitive sensor module for filtration, amplification, analysis and processing. Then the amplified signal is sent to the Ethernet media conversion module.

Step 4: The Ethernet media conversion module is used to facilitate the amplified electrical media conversion into a fiberoptic-compatible optical signal, or light signal. Then this optical signal is sent to the control room/local substation via redundant fiberoptic cable.



Step 5: The fiberoptic cable (FOC) connects the field optical access point (OAP) to the LER communication control switch network. The ICS communication cabinet consists of multiple control switches, which convert the received optical signal from a field optical access point into the corresponding Ethernet-compatible electrical signal.

Step 6: The ICS communication cabinet is connected to the DCS via redundant Ethernet cable copper, twisted pair, Cat. RJ45 or 8P8C connector. The DCS receives the electrical signal from the field instrument for processing and sends the field process-variable information to the display module. The display module presents the process-variable information to the engineering workstation (EWS) or operator workstation (OWS) for further action.

Step 7: After analysis of the process-value information from the field optical instrument, the appropriate control action signal is generated by the DCS to control the process in the plant. The control signal is sent to the ICS communication cabinet via Ethernet cable.

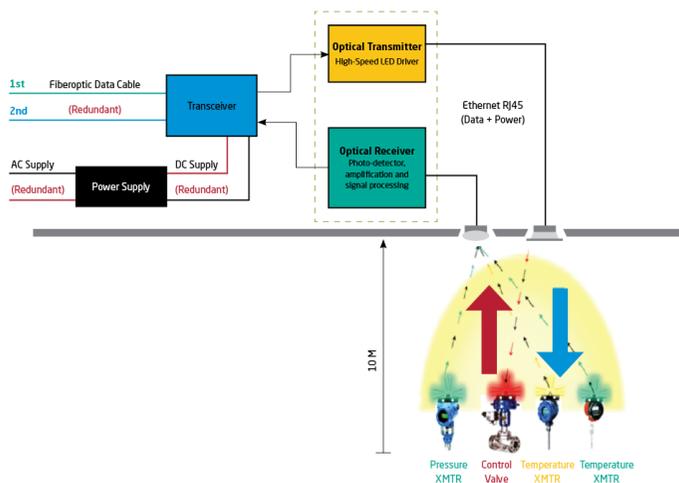
Step 8: The ICS communication cabinet contains multiple control network switches to facilitate the media conversion from Ethernet to fiberoptic-compatible. The fiberoptic cable carrying the corresponding control signal in the form of an optical beam, or light beam, is sent to the field optical access point followed by fiberoptic/Ethernet media conversion.

Step 9: The received control signal from the DCS to the fiberoptic/Ethernet media conversion, located in the field and/or module, converts the optical signal media into an Ethernet-compatible corresponding electrical signal. This electrical signal carries control-action information to the OAP.

Step 10: The OAP receives the control signal from the DCS located in the LCR/SUB via redundant fiberoptic/Ethernet cable. This electrical signal is further encoded in a certain frequency of the switching signal. This switching signal is further sent to the high-speed LED light driver.

Step 11: The high-speed LED light driver regulates the intensity modulation of the LED lights, which is not observed by human eye, as per the received encoded switching signal. The optical access point emits the control signal in the form of an optical signal, which is further received by the control device, such as a valve.

Step 12: The received optical signal carries action information of the final control element and is further converted into an electrical or pneumatic or hydraulic signal to control the process.



References

1. <http://www.warse.org/pdfs/2014/icetetssp25.pdf>
2. <http://www.onlinejournal.in/IJIRV2I6/006.pd>
3. <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6685753>
4. www.oledcomm.com
5. http://ijariie.com/AdminUploadPdf/Review_Paper_on_Li_Fi_Light_Fidelity_ijarii_e2056.pdf
6. <http://www.lifi-lab.com/lifi/lifi-enhanced-mode-future-wireless-communicationreview.html>
7. <https://www.ijsr.net/archive/v5i9/26051603.pdf>
8. <https://www.ijsr.net/archive/v4i12/NOV151778.pdf>
9. http://ijariie.com/AdminUploadPdf/Review_Paper_on_Li_Fi_Light_Fidelity_ijarii_e2056.pdf
10. http://www.ijarcsse.com/docs/papers/Volume_5/6_June2015/V5I6-0175.pdf
11. <http://www.ijcta.com/documents/volumes/vol5issue1/ijcta2014050121.pdf>
12. http://www.academia.edu/6996573/CSE_Study_Paper_on_LiFi_Technology_The_latest_technology_in_wireless
13. <http://www.ijsrp.org/research-paper-0416/ijsrp-p5275.pdf>
14. http://www.academia.edu/6770592/Light_Fidelity_LI-FI_-A_Comprehensive_Study_
15. <https://www.ijedr.org/papers/IJEDR1601019.pdf>
16. <http://www.ijser.org/researchpaper/A-Study-on-LiFi-Light-Fidelity-Technology.pdf>
17. https://www.researchgate.net/publication/279530585_LiFi_Technology_Data_Transmission_through_Visible_Light
18. <http://www.internationaljournalsrg.org/IJECE/2015/Volume2-Issue3/IJECEV2I3P107.pdf>
19. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.666.5679&rep=rep1&type=pdf>
20. http://www.ijarcsse.com/docs/papers/Volume_3/11_November2013/V3I11-0434.pdf
21. https://www.thinkmind.org/download.php?articleid=icwmc_2015_1_10_20026

SECURED DATA TRANSMISSION IN MOBILE AD-HOC NETWORKS USING MULTILEVEL ENCRYPTION TECHNIQUES

Mr.P.Daniel Sundarraj

Department of Computer Science and Application, K.M.G. College of Arts and Science, Gudiyattam

Dr.K.Arulanandam

Department of Computer Application, Government Thirumagal Mills College, Gudiyattam

Abstract

While sending a secret message from source to destination over a wireless network, keeping it safe and secure is a critical issue. An ad-hoc network is a set of wireless nodes and there is no any centralized administration in it. Wireless Ad-hoc networks are configured and organized in self mode. Each node in this type of network is set up using a wireless transmitter and receiver. It allows to transmit data with other nodes in its communication range only. While transmitting data, the nodes communicate using a common physical media. They send and receive signals using the same frequency band, and follow the same hopping method. If the receiving node is not inside the transmission range, the source node uses the other nodes to transmit the messages hop by hop. In order to send a message to another node that is out of its frequency range, it is required to get the help of other nodes in the network for effective data transfer. This is called as multi-hop communication. Hence, every node must act both as a host and as a router at the same time. Wireless Mobile networks are usually attacked by many sources, such as information and physical security attacks. Making wireless ad-hoc networks safe is very difficult for many reasons such as: poor quality of communication channels and nodes, poor infrastructure, frequently changing topology and technology. The wireless channel can be very well and easily accessible by all the network users and malicious attackers. A malicious attacker can easily break network operations by not following the specifications of network protocol. The nodes in the network can easily move in any direction and organize themselves arbitrarily. The nodes can join or leave the network at any time. Because of frequently varying techniques in the network topology, there is a complexity to routing among the various mobile nodes in a safe way. In this paper, multi-level encryption methods are suggested which help us to send the secret message in a more secured way over a wireless network.

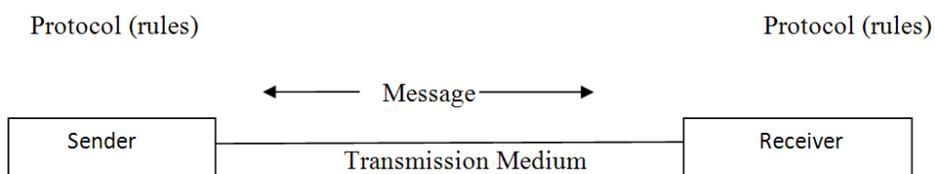
Introduction

Data Communication

Sharing information with others through a communication channel is called data communication. This sharing of information can be local or remote. The local communication occurs face to face, while the remote communication occurs over distance. Computers are very useful devices to exchange information over a network.

Computer Network

A computer network is a group of computers which are interconnected through communication channels. A data communication system has the following important components.



1. **Message:** The data to be transmitted from a sender to a receiver is called Message.
2. **Sender:** Sender is a the source which sends the data to the receiver.
3. **Receiver:** Receiver is a the source which receives the data which is sent by the sender.

4. **Transmission Medium:** The transmission medium is the physical path by which the data are transmitted from a sender to a receiver.
5. **Protocol:** It is a set of rules to be followed while transmitting data from a sender to a receiver through a communication channel.

Network Security

While transmitting data from a sender to a receiver over a network, the data should be protected from unauthorized users. Hence proper principles and methods are to be framed to protect the data which is sent over the network.

Cryptography

Cryptography is a technique which converts an intelligible message into an unintelligible message that could not be understood by others. This process is very essential to send the data from a sender to a receiver and it helps us to protect our data being sent. Since the data travels in an unintelligible format over a network, any unauthorized person in the middle cannot understand it and hence he may not damage or corrupt the data. After the transmitted message received the destination, the message is again converted into its original intelligible format.

The following technical terms are used in cryptography technique.

Plaintext: It is the original intelligible message to be sent from a source to a destination.

Cipher text: The message after converted into unintelligible format is called cipher text.

Cipher: It is an algorithm used for transforming an intelligible message into an unintelligible format.

Key: A specific key is used by the algorithm (Cipher) and it is only known to the sender or receiver.

Encipher (Encoding): It is the process of converting a plaintext to cipher text using a cipher (algorithm) and a key.

Decipher (Decoding): It is the process of converting a cipher text back to its original intelligible format (plaintext format).

Cryptanalysis: It is the study of principles and methods of transforming a cipher text into a plaintext without knowledge of the key. This is called as "code breaking" used by the intruders.

Cryptology: It is both the cryptography and cryptanalysis.

Basic Idea

The Proposed Work has the Following Initiatives

1. For exchanging session key, public key and private keys are used.
2. Certificates will be used to attach these public and private keys to the nodes.
3. Certificates of source and destination are attached with Sending and Reply messages.
4. We propose to use the concept of asymmetric cryptography for exchange of session key
5. We propose to issue Certificates to all participating nodes.

Following symbols are used in the proposed model.

- Source (SO)
- Destination (DE)
- Session key (SK)
- Encrypted Session Key (ESK).

Issuing of Certificates

Let $V = \{V1, V2, V3, \dots, Vn\}$ are different users and $IDE = \{IDE1, IDE2, IDE3, \dots, IDEn\}$ be the identity of users in the wireless mobile ad hoc network and these are unique. Every user V_i has a unique identity IDE_i , which is given to all the other users in the network.

User V_i is assigned an identification number IDE_i and given a signature G_i for IDE_i along with its public and private keys.

$$G_i = (IDE_i^{KE_s}) \bmod m.$$

Issuing of certificates is done by making the users registered and they are given their G_i ($j = 1 \dots m$), public and private keys

Functioning

The Source creates the message, attaches its certificate with keys and sends it to destination. The intermediate nodes resend the message according the protocol. After receiving this message, the destination node verifies the certificate of source and authorized it. The Session key is decrypted at destination first using its private key and then decrypted with the public key.

$$K_{e1} = E_{k_{bd}}(K_s)$$

$$K_{ke} = E_{k_{as}}(K_{e1}).$$

The message is encrypted using session key K_{ke} . After receiving, the source confirms the authenticity of destination from its certificate, decrypts the session key first through its private key and then using public key of destination as $K_{ke1} = D_{k_{bs}}(K_e)$

$K_{ks} = D_{k_{ad}}(K_{e1})$ respectively. Finally session key is used for secure data exchange.

Data Security Design

Data Security is implemented in Two Methods,

1. Data Encryption
2. Data Decryption

Data Security using an Algorithm

- Encrypted data packet is sent to the destination using the route found.
- Finally the destination decrypts the data packet and sends its acknowledgement.

Input Parameters and Output Parameters

Input Parameters:

- At the source, encrypted data packets are sent with destination address and the route request.

Output Parameters:

- At the destination, it receives an acknowledgement with efficient and reliable packet transmission.

Algorithm for Key Generation:

Step 1: Take two prime numbers a and b in such a way that a is not equal to b .

Step 2: Find $m = a \times b$

Step 3: Find $\phi(n) = (a-1)(b-1)$

Step 4: Choose an integer e such that

$$\gcd(\phi(n), e) = 1; 1 < e < \phi(n)$$

Step 5: Find c such that $c = e^{-1} \bmod \phi(n)$

Step 6: Public key $K_{kU} = \{e, m\}$

Step 7: Private key $K_{kR} = \{f, p\}$

Encryption

- The plain text R ($R < m$) is encrypted to cipher text using public key e.
- $G = R^e \pmod{m}$

Decryption

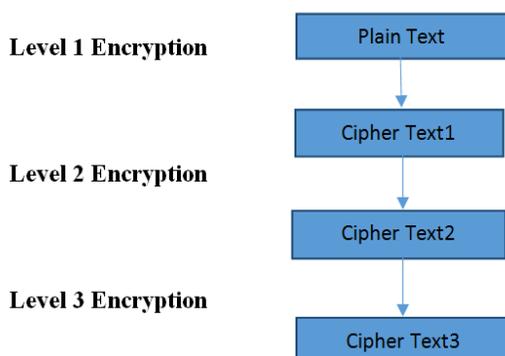
- The cipher text C is decrypted to plain text using private key d.
- $S = C^d \pmod{m}$

Steps for Description

1. If the user wants to send data, the following steps he/she has to do.
2. Get the identifier destination and the encrypted data which is to be transferred.
3. A buffer is to be initialized with the encrypted data to be transferred.
4. A Request Zone is to be set up.
5. A Route Request packet which is having the information about the source and the
6. Destination identifiers, and the Request Zone information will be set up.
7. We need to send the Route Request to its neighbors.
8. An appropriate timer will be set for receiving Route Reply.
9. If the node which receives a packet, we need to identify the type of the packet received.
10. Based on the type of packet received, we need to do following processes.
11. Route Request Process is to be carried out.
12. Process Route Reply is to be carried out.
13. Data Packet is to be carried out.
14. Decryption is to be carried out.
15. Acknowledgement is to be carried out.
16. Route Disconnect is to be carried out.
17. Route Disconnect reply is to be carried out.
18. Timer Run Out is to be carried out.
19. Multi-Level encryption techniques can be used to send the data packets over the wireless adhoc network which improves the security more and more.

Multi-Level Encryption

Multi-Level encryption is a technique of encrypting a plain text more than one time using different keys. Following is the Multi-Level encryption which is accomplished in three levels.



Level 1 Encryption

In the first level of encryption, the Plain Text is encrypted into Cipher Text1.

Level 2 Encryption

In the second level of encryption, the Cipher Text1 is encrypted and it transforms the message into Cipher Text2.

Level 3 Encryption

In the third level of encryption, the Cipher Text2 is encrypted and it transforms the message into Cipher Text3.

Conclusion

After the message is encrypted in multiple levels, it must be decrypted by using decryption algorithms to get the original message. If the message being sent is encrypted in different levels like this, the hackers in the middle may not be easily able to steal the message and damage it. When this method is adopted, time synchronization process is very much essential. To do time synchronization, we need to construct encryption algorithms in such a way that it should not take much time for encryption process.

References

Here we have given various references from which we collected our problem and several others that supported us to design the solution for our problem. These references include either books, papers published through some standards and several websites links with URL's:

1. Kejun Liu, Jing Deng, Member, IEEE, Pramod K. Varshney, Fellow, IEEE, and Kashyap Balakrishnan, Member, IEEE
2. L. Buttyan and J.-P. Hubaux, "Stimulating Cooperation in Self-Organizing Mobile Ad Hoc Networks," ACM/Kluwer Mobile Networks and Applications, vol. 8, no. 5, 2003.
3. K. Balakrishnan is with the Security Services Group, Deloitte and Touche LLP, 1750 Tysons Boulevard, Suite 800, McLean, VA 22102. E-mail: kbalakrishnan@deloitte.com.
4. S. Marti, T. Giuli, K. Lai, and M. Baker, "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks," Proc. MobiCom, Aug. 2000.
5. V.-N. Padmanabhan and D.-R. Simon, "Secure Traceroute to Detect Faulty or Malicious Routing," SIGCOMM Computer Comm. Rev., vol. 33, no. 1, Jan. 2003.
6. V. Srinivasan, P. Nuggehalli, C.F. Chiasserini, and R.R. Rao, "Cooperation in Wireless Ad Hoc Networks," Proc. INFOCOM, Mar.-Apr. 2003.
7. D. Johnson, D. Maltz, Y.C. Hu, and J. Jetcheva, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)," Internet draft, Feb. 2002.

ROLE OF GREEN TECHNOLOGY IN THE FUTURE

D.Saranya

*Assistant Professor, Department of Computer Science and Applications
D.K.M College for Women, Vellore, Tamilnadu, India*

R.Shobana

*Assistant Professor, Department Of Computer Science and Applications
D.K.M College for Women, Vellore, Tamilnadu, India*

Abstract

Computers and Computing devices play a vital role in everyone's life because Computers made our life easy and saves much time and efforts. At the same time, we should be aware of harmful impacts of the use of a computer on the environment. Most of the carbon dioxide (CO₂) emissions are produced through the heat generated by the computing devices. These harmful emissions cause TOXICITY and POLLUTION and affect the environment. In this research paper, we discuss how to overcome these environmental issues by using "GREEN COMPUTING". Green Computing is a technology which aims to reduce the consumption of energy in the computing resources efficiently so that it will not create any impact on the environment. The primary goal of green computing is about minimize CO₂ emissions and hazardous chemicals like CFC's and maximize computing performance to save the environment.

Keywords: Computing devices, Carbondioxide emissions, Toxity, Pollution, Green Computing, CFC's.

Introduction

Green Technology can reduce energy consumption. The temperature of the global world is increasing very quickly. There are many factors, but computers/electronics industry causes over the emission of green house gas and uses much energy consumption which is the cause of current global warming. The energy consumption may be reduced by the introduction of green computing. We can prepare and manufacture such devices which take low energy, give out low heat and gases. Air streams, weather, medicine, transportation, agriculture uses machines which take much amount of power, money and consumption of energy. There are three main areas which affect our daily life, air which we



(Fig:1 Green Computing)
and release low amount of heat.

breathe, water which we drink and food which we eat and the soil on which we live. The data centres use a huge amount of power/energy and release a lot of amount of heat and gases. In our daily life, we use ACs. Refrigerators, inverters, UPS and computers. These items take a huge amount of energy and evolve heat and gases. These gases are very harmful our lives. It has seen that AC and refrigerators release CFC type gases. The battery of inverters also releases dangerous chemicals like lead. It causes lungs type diseases like cancers, asthma. A great amount of heat destroys green house gases like CO₂ which create global warming. A huge amount of heat creates floods, melting of glaciers, drought and increase the temperature of the earth. Many companies are trying to resolve these problems. Companies are trying to establish such devices which can take a low consumption of energy

History of Green Computing

The Green computing came into existence among the drive of work Star program in 1992 by U.S environmental protection author. "Energy Star is a kind of label awarded to computers and other electronics products".

Need for Green Computing

The following points would clear why should a company promote green, or energy efficient computing.

- **Climate Change**

First and foremost, research shows that CO₂ and other emissions are causing global climate and environmental damage. Preserving the planet is a valid goal because it aims to protect existence. Planets like ours, which supports life are very rare. None of the planets in our solar system, or in nearby star systems have m-class planets as we know them.

- **Savings**

Green computing can lead to serious cost savings over time. Reductions in energy costs from servers, cooling, and lighting are generating severe savings for many corporations.

- **Reliability of Power**

As energy demands in the world go up, energy supply is declining or flat. Energy efficient systems helps ensure healthy power.

- **Computing**

Computing Power Consumption has Reached a Critical Point: Data centre's have run out of usable power and cooling owing to high densities.

Current Trends in Green Computing

- **E-Waste Recycling**

Discard used or unwanted electronic equipment in a convenient and environmentally responsible manner. Computers have toxin metals and pollutants that can emit harmful emissions into the environment. Recycling computing equipment such as lead and mercury enables to replace otherwise would have manufactured. The reuse of such equipment allows saving energy and reducing impact on environment.

- **Reduce Paper Consumption**

There are many simple yet effective ways to reduce paper consumption. While using e-mails the electronic archiving uses the track changes feature in the electronic documents rather than using the redline corrections on the paper. While printing any document be sure that both the sides of the paper are utilize. Use of smaller fonts and most importantly selectively printing of required pages.

- **Reduce Power Consumption**

Being energy conscious can help to save many natural resources, water, improve the quality of air, reduce the mercury emissions. Turn off the computers when there's no use for extended period. The power management features should be turn on while there is a short duration of work to be done. With the help of power management techniques, the monitors and computers enter the low power states when they are unused. Power consumption can also be controlled by the following simple tactics like turning the monitor off if the system is going to be idle for a certain amount of time. Also, use of LED and LCD monitors is encouraged over the CRTs as it helps to reduce a lot of power consumption. The hard disks should be turn off when the system is going to be idle; it can be reviewed back by first setting the off time depending upon the usage. Switching to the hibernate mode is another way to save power as this type allows to shut everything down as before shutting down it writes the information in the hard drive whereas in the standby mode the memory is not shut down which consumes power even if the system is idle.

- **Data Center Consolidation & Optimization**

Currently, much of the emphasis of Green Computing area is on Data Centers, as the Data Centers are known for their energy hunger and wasteful energy consumptions.

With the purpose of reducing energy consumption in Data Centers it is worthwhile to concentrate on following:

Information Systems - efficient and right set information systems for business needs are a key in building Green Data Centers. As per green computing best practices efficient servers, storage devices, networking equipment and power supply selection play a key role in the design of information systems.

Cooling Systems – it is suggested by the researchers that at the initial stage of the design process for data centre cooling systems, it is significant to consider both current and future requirements and design the cooling system in such a way, so it is expandable as needs for cooling dictates.

The Standardized environment for equipment is must for Data Centre Air Management and Cooling System. Consider initial and future loads, when designing & Selecting data center electrical system equipment.

- **Virtualization**

One of the foremost trends of green computing is virtualization of computer resources. Virtualization means abstraction of computer resources such as running two or more logical computer systems on a single set of physical hardware. Virtualization helps to achieve a saving of space, saving resources and the environment. Virtualization runs fewer system at higher levels of use. Virtualization allows full consumption of computer resources and benefits in

- Reduction of total amount of hardware.
- Power off Idle Virtual Server to save resources and energy.
- Reduction of total space, air and rent requirements which ultimately reduces the cost

Approaches to Green Computing

- Algorithmic Efficiency
- Power Management
- Video Card
- Display
- Material Recycling
- E-waste management.

Benefits of Green Computing

The following are the enlisted benefits of green computing.

- Cost.
- Efficiency & Improved Performance
- Remote Access
- Disaster Relief
- Scalability
- Reliability and Fault Tolerance
- Ease of Use
- Increased Storage
- Mobility

Security Issues in Green Computing

There are four issues raise while discussing Security in Green Computing.

- Data issues
- Privacy issues
- Security issues
- Infected Applications

Solutions to Security Issues in Green Computing

- Control the Consumer access devices.
- Monitor the Data Access.
- Share Demanded Record and Verify the data's deletion.
- Security Checks events.

Efforts of Green Computing

- **Use Energy Star labeled products:** All the pipeline principal labelled products are manufactured bury maintenance significance mind the term Green Computing again its aspects. These products are manufacture on the suggestion of less intelligence consumption. These devices are programed to power-down to a low power name or when they are not in use. Whence we have to make use of "Energy Star" labelled, monitors, laptops, desktops, printers and other computing piece of equipment.
- **Turn wipe out your computer:** As the previously used figures stated that PC's also its peripherals consume further power and resultant is the great symbol of CO2 emission. So we have to alimony it in our mind and never hesitate to turn off our personal computers when they are not prominence use.
- **Hibernate our computer:** This contour allows us to shut everything secluded. When we are not using our PC considering a short period we count on to hibernate irrefutably. It saves the electricity when the computer is not in use.
- **Turn down guide brightness:** Electricity consumption plays the main role in CO2 emission. If we use our PC at a high brightness, it always put more electricity than using at a usual clarity.

Accordingly, we should always turn down our PC's brightness to save electricity.

- **Use LCD moderately than CRT monitors:** The use of new technologies can play a vital role in pauperize understanding consumption. LCD is the less power consumption device then CRT (Cathode stream cylinder). So if we have to conserve our environment from the effect of CO2 emission, we have to use LCDs rather than CRTs.
- **Abolish informal Disposing:** Computer further its components use toxic chemicals when manufactured and when we gravy informal disposing they put malicious impacts on our environment. In consequence, to reduce these low impacts, we trust to avail formal disposing of.
- **Allow a power plan:** Set a zippy resourcefulness plan to save electricity. As if our computer consumes more electricity; it produce added harmful impacts on our atmosphere.
- **Avoid using screen saver:** smuggle savers are also consumed electricity rolled when a computer is not pulling gain. Screensaver responsibility stands for a graphic, text etc.

Conclusion

This research paper shows the need for Green Computing. We Should Understand the importance of Green Computing, as shown in research paper necessary steps should be taken for a healthy environment. If not then we all will suffer from air pollution, water pollution, soil pollution etc. So with a

little sense of understanding about the importance of Green computing, we should take the steps to "Save Our Environment".

References

1. <http://www.wikipedia.org/>
2. <http://energystar.gov/>
3. <http://thefutureofthings.com/articles /1003 /green-computing.html>
4. www2.hp.com/progra m/care pack/pdf
5. <http://www.wikipedia.org/>
6. <http://searchdatacenter.techtarget.com/ definition/green-computing>
7. Google green computing report
8. <http://www.greencomputing.co.in/>
9. Energy-aware scheduling: <http://domino.research.ibm.com/library/cyberdig.nsf/papers/>
10. <http://timesofindia.indiatimes.com /topic/Green-computing>
11. D. Gahlaut, "Green Computing: Barriers vs Benefits," [Online]. Available: <https://medium.com/@DpanshuGahlaut/the-ultimate-guide-to-green-computing-73e30ba2a485>. [Accessed 03 05 2015].
12. L. Curtis, "Environmental Sustainable Infrastructure Design," The Architecture Journal.
13. M. "Information Systems Innovation for Environmental Sustainability," MIS Quarterly, vol. 34, no. 1, pp. 1-21, March 2010.
14. S. Murugesan, "Harnessing Green IT: Principles and Practices," IEEE Computer Society, pp. 24
15. Green Computing-Embrace a Secure Future International Journal of Computer Applications (0975-8887) Volume 10-N.4, November 2010.
16. GREEN COMPUTING (Internet Sources: Wikipedia; Network World; Environmental Leader; Computer Weekly; Scottish Environment Protection Agency; IT web)
17. Green Computing saves Green- International Journal of Advanced Computer and Mathematical Sciences, Vol 1, Issue 1, Dec 2010.
18. Green Computing from current to future trends-World Academy of Science, Engineering and Technology 63 2012.

CLUSTERING METHODS AND THEIR CLASSIFICATION ON IMAGES

K.Priya

Research Scholar, Bharathiar University, Coimbatore

Abstract

Clustering is the process of organizing objects into groups whose members are similar in some way. A cluster is therefore a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other clusters. Clustering methods are used in various applications in the real world. This paper covers about classification of clustering methods on deferent images.

Keywords: Clustering, images, object, algorithms, classification

Introduction

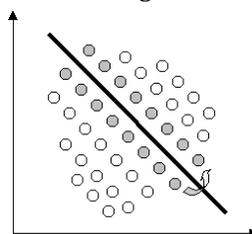
Clustering can be considered the most important unsupervised learning problem; so, as every other problem of this kind, it deals with finding a structure in a collection of unlabeled data. The goal of clustering is to determine the essential grouping in a set of unlabelled data. But how to decide what constitutes a good clustering? It can be shown that there is no absolute "best" standard which would be independent of the final aim of the clustering. Accordingly, it is the user which must supply this standard, in such a way that the result of the clustering will suit their needs. For instance, it could be concerned in finding representatives for homogeneous groups (**data reduction**), in finding "natural clusters" and describe their unknown properties (**"natural" data types**), in finding useful and suitable groupings (**"useful" data classes**) or in finding unusual data objects (**outlier detection**).

Clustering Methods

- Exclusive Clustering
- Density-based methods
- Grid-based methods
- Partitioning methods
- Hierarchical Clustering
- Model-based methods

Exclusive Clustering

In the first case data are grouped in an exclusive way, so that if a certain datum belongs to a definite cluster then it could not be included in another cluster. A simple example of that is shown in the figure below, where the separation of points is achieved by a straight line on a bi-dimensional plane. On the contrary the second type, the overlapping clustering, uses fuzzy sets to cluster data, so that each point may belong to two or more clusters with different degrees of membership.



Density-Based Methods

In density-based clustering clusters are defined as areas of higher density than the remainder of the data set. Objects in these sparse areas - that are required to separate clusters - are usually considered to be noise and border points.

Density Reachability - A point "p" is said to be density reachable from a point "q" if point "p" is within ϵ distance from point "q" and "q" has sufficient number of points in its neighbors which are within distance ϵ .

Density Connectivity - A point "p" and "q" are said to be density connected if there exist a point "r" which has sufficient number of points in its neighbors and both the points "p" and "q" are within the ϵ distance. This is chaining process. So, if "q" is neighbor of "r", "r" is neighbor of "s", "s" is neighbor of "t" which in turn is neighbor of "p" implies that "q" is neighbor of "p".

Advantages

- Does not require a-priori specification of number of clusters.
- Able to identify noise data while clustering.
- DBSCAN algorithm is able to find arbitrarily size and arbitrarily shaped clusters.

Disadvantages

- DBSCAN algorithm fails in case of varying density clusters
- Fails in case of neck type of dataset.

Grid-based Methods

Grid-based methods quantize the object space into a finite number of cells that form a grid structure. All the clustering operations are performed on the grid structure. The advantage of this approach is fast processing time, STING, CLIQUE, and Wave-Cluster are examples of grid-based clustering algorithms. The grid-clustering algorithm is the most important type in the hierarchical clustering algorithm. The grid based clustering approach considers cells rather than data points. This is because of its nature grid-based clustering algorithms are generally more computationally efficient among all types of clustering algorithms. In fact, most of the grid-clustering algorithms achieve a time complexity of $O(n)$ where n is the number of data objects. It allows all clustering operations to perform in a gridded data space. Grid-based clustering where the data space is quantized into finite number of cells which form the grid structure and perform clustering on the grids. Grid based clustering maps the infinite number of data records in data streams to finite numbers of grids. Grid based clustering is the fastest processing time that typically depends on the size of the grid instead of the data. Grid-based methods are highly popular compared to the other conventional models due to their computational efficiency. The main variation between grid-based and other clustering methods is as follows. In grid based clustering all the clustering operations are performed on the segmented data space, rather than the original data objects. Then any topological neighbour search is used to group the points of the closer grids. The grid-based clustering methods face the following challenges. First, it has to determine an appropriate size of the grid structure. If the grid size is too large, two or more clusters may be merged into single one. When the grid size is very small, a cluster may be divided into several sub-clusters. Therefore, finding the suitable size of grid is a challenging issue in grid clustering methods. The second problem is with the data of clusters with variable densities and arbitrary shapes in case of which a global density threshold cannot result the clusters with less densities. This is known as the problem of locality of cluster. The third one is the selection of merging condition to form efficient clusters. Considering these issues, various grid based algorithms have been proposed. Thus these algorithms have a fast processing time, because they go through the data set once to compute the statistical values for the grids and the performance of clustering depends only on the size of the grids which is usually

much less than the data objects. Grid-based methods are widely used in today's various fields applications such as Pattern Recognition, Spatial Data Analysis, Image Processing, WWW (World Wide Web) and others.

Partitioning Methods

Perhaps the most popular class of clustering algorithms is the combinatorial optimization algorithms a.k.a. iterative relocation algorithms. These algorithms minimize a given clustering criterion by iteratively relocating data points between clusters until a (locally) optimal partition is attained. In a basic iterative algorithm, such as K-means- or K-medoids, convergence is local and the globally optimal solution can not be guaranteed. Because the number of data points in any data set is always finite and, thereby, also the number of distinct partitions is finite, the problem of local minima could be avoided by using exhaustive search methods. However, this is truth only in theory, since finding the globally optimal partition is known to be NP-hard problem and exhaustive methods are not useful in practice. The number of different partitions for n observations into K groups is a Stirling number of the second kind, which is given by $S(K, n) = \frac{1}{K!} \sum_{i=0}^{K-1} (-1)^{K-i} \binom{K}{i} i^n$. This shows that enumeration of all possible partitions is impossible for even relatively small problems. The problem is even more demanding when additionally the number of clusters is unknown. Then the number of different combinations is the sum of the Stirling numbers of the second kind: $\sum_{i=1}^{K_{max}} S(i, n)$, where K_{max} is the maximum number of cluster and it is obvious that $K_{max} \leq n$. The fact is that exhaustive search methods are far too time consuming even with modern computing systems. Moreover, it seems be an infinite race between computer power and amount of data, which both have increased constantly during the last years. Therefore, more practical approach than exhaustive search is the iterative optimization.

Hierarchical Clustering

A hierarchical clustering algorithm is based on the union between the two nearest clusters. The beginning condition is realized by setting every datum as a cluster. Connectivity based clustering, also known as hierarchical clustering, is based on the core idea of objects being more related to nearby objects than to objects farther away. As such, these algorithms connect "objects" to form "clusters" based on their distance. A cluster can be described largely by the maximum distance needed to connect parts of the cluster. At different distances, different clusters will form, which can be represented using a dendrogram, which explains where the common name "hierarchical clustering" comes from: these algorithms do not provide a single partitioning of the data set, but instead provide an extensive hierarchy of clusters that merge with each other at certain distances.

Model Based Clustering Methods

This algorithm is based on hypothesizing a model for every cluster to find best fit of the data according to the mathematical model. It can automatically determine the number of cluster based on standard statics. The method may locate clusters by constructing a density function that reflects the spatial distribution of the data points. This method also provides a way to automatically determine the number of clusters based on standard statistics taking outlier. It therefore yields robust clustering method. STASTICAL approach and COBWEB are examples of model based clustering methods.

Most used Clustering Algorithms are

- K-means
- Fuzzy C-means
- Hierarchical clustering

K Means Clustering

The K-means clustering is similar or simpler and easier in computation of an efficient K-means clustering algorithm it is the simplest unsupervised learning algorithm that solve the well-known clustering problems. It's the K-means algorithm is an unsupervised clustering algorithm that classified in the input data points into multiple classes based on their intrinsic distance from other dataset points of this cluster. It assume that the data features from a vector space and tries to find natural clustering.

Fuzzy C-Means Clustering

Clustering, a major area of study in the scope of unsupervised learning, deals with recognizing meaningful groups of similar items. Among all clustering algorithms, the Fuzzy C-means (FCM) clustering algorithm FCM calculates membership values for each data point to all clusters it usually generates a more consistence and trust-worthy results.

Hierarchical Clustering

The hierarchical agglomerative clustering methods are most commonly used. The construction of an hierarchical agglomerative classification can be achieved by the following general algorithm.1. Find the 2 closest objects and merge them into a cluster 2. Find and merge the next two closest points, where a point is either an individual object or a cluster of objects. If more than one cluster remains, return to step 2 Two main approaches:

Where the two approaches will be done on it. where

- Agglomerative approach
- Divisive approach

Individual methods are characterized by the definition used for identification of the closest pair of points, and by the means used to describe the new cluster when two clusters are merged.

Applications

Some of the practical applications of image are :

- Content-based image retrieval
- Machine vision
- Medical imaging, including volume rendered images from computed tomography and magnetic resonance imaging.
 - Locate tumors and other pathologies
 - Measure tissue volumes
 - Diagnosis, study of anatomical structure
 - Surgery planning
 - Virtual surgery simulation
 - Intra-surgery navigation
- Object detection
 - Pedestrian detection
 - Face detection
 - Brake light detection
 - Locate objects in satellite images (roads, forests, crops, etc.)
- Recognition Tasks
 - Face recognition
 - Fingerprint recognition

- Iris recognition
- Traffic control systems
- Video surveillance

Several general-purpose algorithms and techniques have been developed for image segmentation. To be useful, these techniques must typically be combined with a domain's specific knowledge in order to effectively solve the domain's segmentation problems.

Conclusion

The goal of clustering is grouping of similar objects. In this paper discussed various clustering methods which are used on images in different way to analyse the part of an image. Also discussed most frequently used algorithms and applications of clustering on various image processing areas. Clustering methods are used to store, organize and integrate the similar kind of objects in various fields in the real life.

References

1. http://www.elet.polimi.it/upload/matteucc/Clustering/tutorial_html/kmeans.html
2. Andrew Moore: "K-means and Hierarchical Clustering - Tutorial Slides"
<http://www-2.cs.cmu.edu/~awm/tutorials/kmeans.html>
3. Gonzalez, R.C.; Woods, R.E. Digital Imaging Processing; Prentice Hall: New York, NY, USA, 2002.
4. https://en.wikipedia.org/wiki/Image_segmentation
5. <https://www.coursera.org/learn/cluster-analysis/lecture/grfVg/5-4-grid-based-clustering-methods>
6. Damodar Reddy Edla and Prasanta K. Jana "A Grid Clustering Algorithm Using Cluster Boundaries" IEEE World Congress on Information and Communication Technologies 2012
7. http://users.jyu.fi/~samiayr/pdf/introtoclustering_report.pdf
8. Prabhdeep Kaur, Shruti Aggrwal , Comparative Study of Clustering Techniques, international journal for advance research in engineering and technology, April 2013.
9. P. Dhanalakshmi & T.Kanimozhi "Automated segmentation of Brain tumor using K.means clustering and its area calculation - IJAREEE" 2013.
10. Google:<http://sites.google.com/site/dataclusteringalgorithms/k-mean-clustering-algorithm>
11. <https://www.slideshare.net/EdurekaIN/applications-of-clustering-in-real-life>

MANAGEMENT OF NETWORK FROM ATTACK-RANSOM WARE

S.Asmathullah

Research Scholar, Department of Computer Science, Islamiah College (Autonomous), Vaniyambadi

V.A.Afsal Basha

Associate professor, Department of Computer Science, Islamiah College (Autonomous), Vaniyambadi

Abstract

Cyber security is the internet and the Internet because of the leak of important data through the Internet and indicates that the term is called cyber security. Computer security covers all processes and processes that protect the information of the secrecy, integrity and information availability of both digital equipment, information and services, unwanted or unauthorized access, alteration or destruction, and transit and rest location. Cyberspaces and its basic infrastructure are weakening a wide range of risk against both physical and cyber threats and risks. Modern cyber actors and states of the states face threats to steal information and money and there are development capabilities for the delivery of damaged, destroying, or threatening delivery of necessary services. A range of traditional crimes is now being set up by cyber space. It includes child transmission and child abuse conspiracy, banking and financial fraud, intellectual property violations, and other crime production and distribution, of which everything is quite human and economic consequences.

Keywords: Data modification, identifying spoofing, man-in-middle attack, sniffer attack, Virus, malware, Ransom ware.

Introduction

Information security is often defined as security or guarantee of information and requires the ability to maintain the authenticity of information. The three common components of information security are confidentiality, integrity and availability and form an essential basis for the general framework of information security. Confidentiality is perhaps one of the most common aspects of information security because any information that is held by the public with the intent to allow only access to authorized persons is reserved. Encryption is often used to maintain information privacy because encryption is one of the most access control methods and Microsoft Bit Locker is an example of access control encryption. Although True Crypt is no longer a safe way to protect data, I will discuss it further in the Cyber Security part of the document. In addition to confidentiality, integrity is an important aspect because the format or original content of the information should not under any circumstances change except with the authorization of authorized individuals. The hashing method is commonly used to ensure integrity and the hash of the source must match the hash of the destination to ensure that the data is free of any changes. Finally, information should be available to all authorized individuals and this component is usually targeted by denial-of-service or DOS attacks. Availability is essential for the continuity of a company in times of disaster, extreme emergency or data breach, so it strongly depends on the ability of a company or an agency to provide services and to constantly back up all data. In the event of a breach or "hack", availability is a top priority as security engineers strive to seal access to confidential data while allowing those who are authorized to access the data. The interrelationship between confidentiality, integrity and availability is essential to establish policies and procedures in information technology.

Common Types of Network Attacks

Data Modification

After an attacker has read your information, the next logical step is to change it. Without the knowledge of an attacker sender or receiver, the data of the packet can be changed. Even if you do not require privacy for all communications, you do not want to change any transit of your messages. For example, if you are exchanging for purchase, you do not want to change the item, quantity or billing information

Identity Spoofing (IP Address Spoofing)

Most networks and operating systems use the computer's IP address to identify the actual object. In some cases, an erroneous IP address is mistakenly allowed - an identity counterfeit. An attacker can also use special programs to create IP packets that appear to originate from valid addresses within a corporate intranet. By accessing a network with a valid IP address, an attacker can modify, redirect, or delete your data. An attacker can also conduct other types of attacks, as described in the following sections.

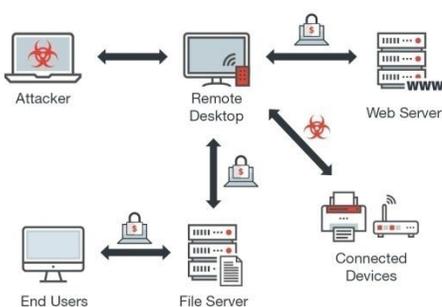
Man-in-the-Middle Attack

As the name suggests, a man-in-the-middle assault occurs when someone between you and the person you are communicating with is actively monitoring, capturing and controlling your communications in a transparent manner. For example, an attacker can divert a data exchange. When the computer communicates on a low level of network layer, the computer may not be able to determine with whom they exchange data. Human attack in the middle is like someone who considers your identity to read your message. People on the other may believe it is you because an attacker can react actively to keep an active exchange and get more information. This attack is capable of damage similar to the app-level attack, which is described later in this section

Sniffer Attack

A sniffer is an application or device that can read, control and capture network data exchanges and read network packets. If the packages are not encrypted, the snakes provide a complete look inside the package. Even encapsulated packages can be opened and read if they are not encrypted and the attacker does not have the key. Using a sniper, the attacker can do the following: Analyze your network and get information that ultimately leads to your network crash or corrupt. Read your contacts?

Ransom ware Attack



The revocation device is a form of malicious software from crypto virology that threatens to print the victim's data or constantly disrupts its access unless a ransom is paid. While some simple silver rations may lock the system in a way that is unwise for a discerning man to change, the more advanced malware uses a method called crypto viral extortion, where it encrypts victim files, makes them inaccessible, and requires a ransom payment to decrypt them. With a strictly enforced crypto viral extortion attack, the recovery of files without the decryption key is an inevitable problem - and it is difficult to track digital

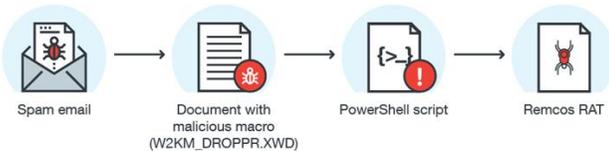
money like Ukash and Bit Coin used for ransoms, making monitoring and analysis of hard-working.

A Case Study in Ransom Attack

Ransom ware is a form of malicious software that locks the files on your computer, encrypts them and requires you to pay to get your files back. Wanna-Decrypt or, or Wanna-Cry, is a kind of ransom that affects Microsoft's Windows operating system. When a system is infected, a popup window will appear asking you to pay to restore all your files within three days of a countdown timer to the left of the window. It adds that if you do not pay within that time, the fee will be doubled and if you do not pay within seven days you will lose the files forever. Payment is only accepted with Bit coin

How does work?

The Securities of the US Emergency Readiness Team (USCRT), the Human Rights Security



Department, could not restrict the freedom of the program, but the software is not available for software. What WannaCry uses is propagated by a number of trademarks - which differs from the game in a computer or computer,

because the system has become disabled for the purpose of searching. The National Security Agency (NSA) agreed to provide information on the situation in the United States. No malware form is the perfect way to completely delete the computer, but it does not matter how you use it. The 2017 magnetic field arrives at 200,000 twenty-two and twenty-five hours of computers of 150 percent

There are a Few Dos and Don'ts When it Comes to Ransom Ware

Do not pay the ransom. It encourages and finances only these aggressors. Even if the ransom is paid, there is no guarantee that you will be able to regain access to your files. Restore all affected files from a known backup. Restoring your files from a backup is the quickest way to regain access to your data. Do not provide personal information when replying to an e-mail, an unsolicited call, a text message or an instant message. Phishing will try to trick employees into installing malware or acquiring information for attacks by claiming to come from IT. Be sure to contact your IT department if you or your colleagues receive suspicious calls. Use reliable antivirus software and a firewall. Maintaining a strong firewall and keeping security software up to date is crucial. It is important to use the reputable company's antivirus software due to all the fake software out there. Use content scanning and filtering on your mail servers. Incoming emails should be scanned for known threats and should block all types of attachments that could pose a threat. Make sure all systems and software are up to date with the relevant patches. Exploit kits hosted on compromised websites are commonly used to spread malware. Regular patching of vulnerable software is needed to help prevent infection. If you travel, notify your IT department in advance, especially if you use the public wireless network. Make sure you are using a reliable virtual private network (VPN) when accessing public Wi-Fi like Norton Wi-Fi Privacy.

In the Event of a Ransom ware Attack

While these practices are effective, it is impossible to completely protect your ransom organization. If you believe you have been the victim of a ransom attack, consider the following steps:

Take a snapshot of your system. Before turning off the system, if possible, try to capture a snapshot of system memory. This will help you later locate the attack vector of the ransom ware, as well as any cryptographic material, which can help with decrypting the data. Turn off the system. To prevent the further ransom ware from spreading and the inevitable damage to data, turn off the system that is believed to be infected. Identify the attack vector. Recall all e-mails suspected of bringing the ransom

attack to prevent further spread of the attack. Block access to the network to all command and control servers identified by ransom ware. Ransom ware is often blocked by data encryption without access to these servers. Notification to the authorities. Consider informing the authorities so they can help with the investigation. Even if law enforcement agencies can help with an investigation, there is also a risk that data can never be recovered. Redemption payments tend to rise when time passes for payment to be made. The involvement of law enforcement agencies could also delay and add a significant cost to the redemption if eventually the user (s) decides to pay.

Solution for Ransom ware

Digital Guardian threat detection and response software offers comprehensive protection against advanced cyber attacks, including redemption. Detects and blocks advanced threats throughout the life cycle of attacks. Our threat detection and response capabilities have successfully detected and contained the WannaCry attack for ALL our advanced threat protection customers.

- Back up your data
- Show hidden file-extensions
- Filter EXEs in email
- Disable files running from AppData/LocalAppData folders
- se the Crypto locker Prevention Kit
- Disable RDP (remote desktop protocol)
- Patch or Update your software
- Disconnect from Wi-Fi or unplug from the network immediately
- Use System Restore to get back to a known-clean state
- Change password frequently and password must be complex.

Conclusion

We shown that their impact can be devastating for small business owners and organizations. Ransom ware is not just a threat to small businesses and the organization has an impact on people. In his FBI public service request report, they ask anyone who has suffered a ransom infection to never pay a ransom because it helps criminals perfect their attacks and pin more victims. The FBI says that paying a ransom does not guarantee that the victim can regain access to their data; indeed, some people or organizations never receive decryption keys after paying a ransom [16 May 2017]. The recommendation that I believe can help small business owners and organizations prevent and defend ransom attacks using Trend Micro Security 10, VIPRE Internet Security Pro Small Office and Kaspersky Internet Security.

References

1. Cohen f.a short course on computer virus ,new York:wiley,1994
2. Hisao,D.et,al Computer Security,1979
3. Denning,D. Information Cyber Security,Wesley,1997
4. omicsonline.org/scholarly/cyber-security-journals-articles-ppts-list.php
5. dhs.gov/cybersecurity-overview
6. ist.psu.edu/students/undergrad/majors/sra/ic

NETWORK INTELLIGENCE FOR WIRELESS DATA TRANSMISSION USING SOCKET PROGRAMMING

M.Dukitha

Assistant Professor MCA, Er.Perumal Manimekalai College of Engineering-Hosur

T.Balasubramanian

*Associate Professor, CS, Sri Vidya Mandir Arts and Science College
Uthangarai, Krishnagiri (Dt),Tamilnadu*

M.Aruna

Second Year MCA, Er.Perumal Manimekalai College of Engineering-Hosur

P.Kokila

Second Year MCA, Er.Perumal Manimekalai College of Engineering-Hosur

Abstract

Wireless communication is one of the major necessary mediums of transmission of data from one device to different devices. During this technology, the data is transmitted through the air while not requiring any cable or wires or different electronic conductors, by victimization magnetism waves like IR, RF, satellite, etc. within the days, the wireless communication technology refers to spread to wireless communication devices and technologies starting from the devices like sensible phones to computers, tabs, laptops, Bluetooth Technology, printers. The data transmission is applied in 2 modes synchronous transfer mode and asynchronous transfer mode. The early wireless communication devices were introduced and also the technology has considerably been developed over intervening and resultant years. Today, the term wireless refers to a spread of devices and technologies starting from sensible phones to laptops, tabs, computers, printers, Bluetooth, etc. During this gone world wherever information transmission plays an important role for communication in any kind, it becomes terribly necessary to use the network intelligence for wireless information transmission with less traffic overhead.

Keywords: *Wireless, Network programming, Socket programming, Wi-Fi, Bluetooth.*

Introduction

Sockets provide an interface for communication between computers by TCP. A client program creates a socket on its end of the communication and attempts to connect that socket to a server. When the connection is made, the server creates a socket object on its end of the communication. The client and server can now communicate by put and get data from the socket [7]. Where one program start the communication called client process or program and other who is waiting for communication to start called the server process or program. In the simplest case, Client program sends the request to the server. Server sends the response. Client then obtains the data from the server and displays it. Complex clients filter and reorganize data, repeatedly obtain changed data, send data to other people , and allows real time chatting, multiplayer gaming [4]. Socket programming can be use in many languages, but they must use for supports network communication, but java is better because it is platform independent, it has exception mechanisms for robust handling of common problems that occur during I/O and networking operations and its threading facilities provide a way to easily implement powerful servers One of java's major strong suits as a programming language for client server application development is its wide range of network support [8].

Whenever we want to communicate with one computer to other computer over network then it is required to software which connects two ports in network layer that is called network programming.

After connecting over network, its play a big role to send and receive data to millions of people over world [6].

This is about a client/server multi-threaded socket class. The developer still responsible to decide if the thread is needs up to that the thread is optional. There are other Socket classes here and other places over the Internet but none of them can provide feedback (event detection) to your application like this one does [2].

Transmission Control Protocol

The transmission control protocol (TCP) communication. This limitation is bad if all you want is a simple console application. So this library doesn't have such a limitations. It also provides threading support automatically for you, which handles connection and disconnection to a peer [2]. TCP socket provides the communication mechanism between two devices particular distance to another device. This is managed by the introduction of a port which is numbered socket on particular devices [5].

Disadvantages of Transmission control protocol

- TCP cannot be use broadcast and multicast connection,
- High cost of cables,
- Any one place cable is damaged entire cable have changed,
- Face many problems in suffering and downloading contents from the net, and
- Particular distance only communicates with TCP.

User Datagram Protocol

User datagram protocol uses alternative protocol for sending data. UDP provides connectionless communication. The connection here would be unreliable. Datagram socket class contains constructor that is used to construct a datagram socket and bind it to the specified port on the local host machine. The client does not establish a connection with the server like in case of UDP. Instead, the client just sends a datagram to the server using send to function which requires the address of the destination as a parameter. Same as , the server not accept a connection from a client. For that , the server calls the function recvfrom, that waits for client data arrive.

Advantages of User Datagram Protocol

Some of the advantages of User Datagram protocol

- Include flexibility
- Cost effectiveness, High speed of data transmission
- Constant connectivity
- Convenience
- Enhanced efficiency
- Easy to data transmission using wireless communication,

Comparison between TCP/IP and UDP

There is some major difference between TCP and UDP protocol in java application developer's kit.

1. TCP is a reliable stream oriented protocol as opposed to UDP which is not reliable and based upon datagram. The actual difference between TCP and UDP is that UDP cannot use for sending important messages. Though there are some reliable protocols built over UDP e.g. TIBCO certified messaging which implement additional checks whether the message is delivered or not and then facilitate re-transmission.

2. Another difference between TCP and UDP in the speed. Since TCP is connection oriented and reliable. It has lots of overhead as compared to UDP, which means TCP is slower than UDP and should not be used for transferring message where speed is critical e.g.
3. Another structural difference between TCP and UDP is that data boundaries are preserved in case of UDP but not in the case of TCP because data is sent as it is as one message in case of UDP but TCP protocol can break and reassemble the data at sending and receiving end.
4. Another key difference between TCP and UDP protocol comes from the fact that TCP is a connection-oriented protocol but UDP is a connectionless protocol. What this mean is, before sending a message a connection is established between sender and receiver in TCP but no connection exists between the sender and receiver in UDP protocol.

Literature Review

The socket programming concept has used in many areas of networks. The socket application developed in this paper is based on the distributed computing concept. This is relevant as to that it allows for resources to be distributed among several node in the network [4]. Socket programming is a suitable approach to develop a not-too-complex networked application. This technique involves creating a socket which is used as a communication channel to transmit and receive messages [5]. We can use socket programming to develop application that is based up on two typical architectures, namely the client-server and peer to peer. Internet and WWW have emerged as global ubiquitous media for communication and changed the way we learn, live, enjoy, communicate, interact, engage, etc [6]. So the library doesn't have limitation. It automatically give threading support for you, which handles the socket connection and disconnection to a peer. It also features some options not yet found in any socket classes that I have seen so far [2]. Socket programming can be done in any language, which supports network communication, but java is better because it is platform independent language, it has exception mechanisms for robust handling of common problems that occur during I/O and networking operations and its threading facilities provide a way to easily implement powerful servers One of java's major strong suits as a programming language for client-server application development is its wide range of network support [8].

Types of Wireless Communication

The different types of wireless communication mainly include,

- Satellite communication,
- Broadcast radio,
- Microwave communication,
- Radio,
- Infrared communication,
- Wi-Fi communication,
- Mobile communication systems.



Figure 1 Types of wireless communication [10]

Satellite Communication

Satellite communication works based on the self-contained wireless communication technology, it is commonly used over all the world. This is permitting the users to linked them all over on the earth. When the signal (a beam of modulated microwave) is pass near the satellite then, satellite amplifies the signal and return back the signal to the receiver antenna, that is located on the earth. Satellite communication contains two main components like the space segment and the ground segment. The ground segment consists of fixed or mobile transmission, reception and ancillary equipment and the space segment, which mainly is the satellite itself.



Figure 2 Satellite communication [10]

Advantages of Satellite Communication

- It is used for mobile and wireless communication applications independent of location.
- It covers wide area of the earth hence entire country or region can be cover with just one satellite.
- It is used for voice, data, video and any other information transmission. Satellite system can be interfaced with internet infrastructure to obtain internet service. It also used for GPS applications in various devices for location determination.

Disadvantages of Satellite Communication

- Satellite has been constructed for years. Moreover satellite design and development requires high cost.
- Satellite once launched, requires to be monitored and controlled on regular periods so that it remains in the orbit.
- Satellite has life which is about 12-15 years. Due to this fact, another launch need to be planned before it becomes un-operational.

Broadcast radio

The very first wireless communication technology is an open radio communication. It used in many ways. Still now it uses in many ways. Hand held multichannel radios allows an user to speak with in short distances, whereas citizen's band and maritime radios offer communication services for sailors. Ham radio enthusiast divide data and function for emergency communication allover disasters with their powerful broadcasting gear, and can even communicate digital information over the radio frequency spectrum.

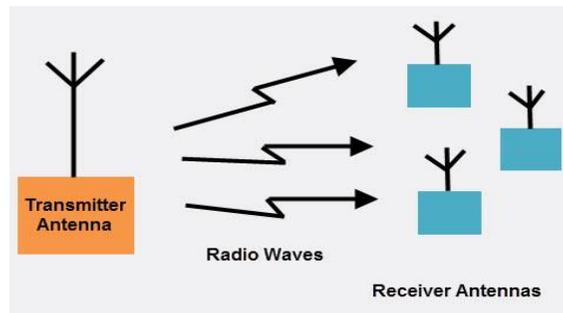


Figure 3 Broadcast radio [10]

Microwave Communication

The Microwave wireless communication is an effective communication technology. This transmission technology uses in radio waves, also used in wavelengths of radio waves are measured in centimeters. In this communication, the data or information can be transfers using two methods. One is satellite method and another one is terrestrial method. The frequency range of the terrestrial system is typically 4GHz-6GHz and with a transmission speed is usually 1Mbps to 10Mbps.



Figure 4 Microwave Communication [10]

Radio Communication

For example, you can take a radio station. When the RJ says you are listening to 92.7 BIG FM, what he really means is that signals are being broadcasted at a frequency of 92.7 megahertz, that successively means the transmitter at the station is periodic at a frequency of 92,700,000 Cycles/second. When you would like to listen to 92.7 BIG FM, all you have to do is tune the radio to just accept that specific frequency and you will receive perfect audio reception.



Figure 5 Radio communication [10]

Advantages of Radio Communication

- Radio is one of the media which covers huge population.
- Radio can be enjoyed at home, in office, while driving car can be enjoyed anywhere.

- You can advertise your product on radio and the rate of advertisement is usually lower than other medium of communication.
- It is used to important information or news can be easily spread on radio.

Disadvantages of Radio Communication

- Only an audio medium of communication.
- During bad weather you cannot listen radio properly. Often unclear and is affected by weather.
- Less and limited radio channels are available compared to other communication medium.

Infrared Communication

Infrared wireless communication transfer information between computers through IR radiation. IR is electromagnetic energy with the wavelength that has maximum wave length up to red light. It is used for security control, TV remote control and short range communications. In the electromagnetic spectrum, IR radiation lies between microwaves and visible light. So, they can be used as a source of communication.

Advantages of Infrared Communication

- The devices are very compact, lightweight and consume low power.
- The technology based devices are easy to use.
- It is more secure compare to RF technology.

Disadvantages of Infrared Communication

- It requires both transmitter and receiver to be inline of sight.
- Devices cannot move around while transmission is in progress.
- Used for very short distance applications.



Figure 6 Infrared Communication [10]

Wi-Fi Communication

Wi-Fi is a low power wireless communication that is used by various electronic devices like smart phones, laptops, etc. In this setup, a router works as a communication hub wirelessly. These networks allow users to connect only within close proximity to a router. Wi-Fi is very common in networking applications which affords portability wirelessly. These networks need to be protected with passwords for the purpose of security, otherwise it will access by others.



Figure 7 Wi-Fi Communication [10]

Advantages of Wi-Fi Communication

- The main advantages of using Wi-Fi technology is the lack of wires. This is a wireless connection that can merge together multiple devices.
- Wi-Fi network is particularly useful in cases where the writing is not possible or even unacceptable. For example it is often used in the halls of conferences and international exhibitions.
- Wi-Fi networks are widely used to connect a variety of devices, not only between themselves but also to the internet. And almost anywhere, not just where the cables are laid.

Disadvantages of Wi-Fi Communication

- Call quality is greatly influenced by the environment, is particularly sensitive to electromagnetic radiation generated by household appliances. The primarily affects the speed of communication.
- Despite the global standardization, many devices from different manufactures are not fully compatible, which in turn affects the speed of communication.
- Really Wi-Fi technology is not perfect and has many flaws that limit its use. However, the benefits of it is much greater. Therefore, every day, this modern technology is increasingly used and becomes a popular.

Mobile Communication System

The advancement of mobile networks is enumerated by generations. Many users communicate across a single frequency band through mobile phones. Cellular and cordless phones are two examples of devices which make use of wireless signals. Typically, cell phones have a larger range of networks to provide coverage. But, Cordless phones have a limited range. Similar to GPS devices, some phones make use of signals from satellites to communicate.

Bluetooth technology

The main function of the Bluetooth technology is that permits you to connect a various electronic devices wirelessly to a system for the transferring of data. Cell phones are connected to hands free earphones, mouse, and wireless keyboard. By using Bluetooth device the information from one device to another device. This technology has various functions and it is used commonly in the wireless communication market.

Advantages of Bluetooth technology

- It is cheap,
- Easy to install,
- It makes connecting to different devices convenient,
- It is wireless,
- It is free to use if the device is installed with it.

Disadvantages of Bluetooth Technology

- It can be hacked into,
- If installed on a cell phone it is prone to receiving cell phone viruses,
- It only allows short range communication between devices,
- It can only connect two devices at once,
- It can lose connection in certain conditions.

Applications of Wireless Communication

- Applications of wireless communication involve security systems,
- Television remote control,
- Wi-Fi,
- Cell phones,
- Wireless power transfer,
- Computer interface devices and various wireless communication based projects.

Conclusion

Developing wireless communication applications is made possible in Java by using sockets, threads, Wi-Fi, satellite, Bluetooth, broadcast radio, radio, infrared and microwave. These wireless technologies are used to how to transmission or communicate with one system to another system. These are explain in this paper.

Typically, programs running on client machines make requests to programs on a server machine. These involve networking services provided by the transport layer. The most widely used transport protocols on the Internet are TCP (Transmission control Protocol) and UDP (User Datagram Protocol). TCP is a connection-oriented protocol providing a reliable flow of data between two computers. It is used by applications such as the World Wide Web, e-mail, IP, and secure shell.

Sockets provide an interface for programming net-works at the transport layer. Using sockets, network communication is very much similar to per-forming file I/O. A socket is an endpoint of a two-way communication link between two programs running on the network. The source and destination IP address, and the port numbers constitute a net-work socket.

References

1. "Socket Programming" [Http://Www.Cs.Rutgers.Edu/~Pxx/Rutgers/Notes/Sockets/Index.Html](http://Www.Cs.Rutgers.Edu/~Pxx/Rutgers/Notes/Sockets/Index.Html)
2. Saranya C, Gomathi L, "Multitasking – Sockets In Data Transmission" August '2015.
3. Maha Sabri Altemem, "Voice Chat Application Using Socket Programming" Vol. 2, Issue 5, August '2014.
4. Sweety Sen1, Sonali Samanta1, "Network Programming In Java Using Socket" Volume 3 Issue 5, May' 2015.
5. Limikalita, "Socket Programming", Vol. 5 (3), 2014, 4802-4807.

6. Ranjeet Kumar, Vinay, "Elements Of Socket Programming In Java" © 2014 IJIRT |,Vo1., Issue 6, ISSN: 2349-6002
7. Prerna Malik, Poonamrawat", Network Programming In Java Using Sockets", 2014 IJIRT, Volume 1, Issue 6, ISSN: 2349-6002.
8. Sandeep Singh, Hemant, "Review Paper Programming In Sockets" © 2015 IJIRT, Volume 1, Issue 12, ISSN: 2349-6002.
9. Dnyaneshwaribhusari, Gauravmokase, Prasad Waghmare. "Wifi Calling Android Phones In Socket Programming", January 2015.
10. "Wireless Communication" [Http://Www.Efxkits.us/Different-Types-Of-Wireless-Communication-Technology](http://Www.Efxkits.us/Different-Types-Of-Wireless-Communication-Technology).

MAJOR ISSUES AND CHALLENGES ARE INVOLVE IN WIRELESS AD HOC NETWORK AND ITS SOLUTION

M.Inbavalli

Associate Professor, MCA, Er.Perumal Manimekalai College of Engineering, Hosur

M.Angelinrosy

Assistant Professor, MCA, Er.Perumal Manimekalai College of Engineering, Hosur

C.Anitha

Project Lead, Infosys, Bangalore

S.Rekha

Second Year MCA, Er.Perumal Manimekalai College of Engineering, Hosur

S.Nivetha

Second Year MCA, Er.Perumal Manimekalai College of Engineering, Hosur

Abstract

Traditionally, military science networks are the sole communication networking application that followed the accidental paradigm. Recently the introduction of latest technologies like Bluetooth and hyperlan area unit serving to change ultimate business mobile accidental network readying outside the military domain. The recent evolutions are generating a revived and growing interest within the analysis and development of mobile accidental network. Mobile accidental network represent complicated distributed systems that comprise wireless mobile nodes which will freely and dynamically self-organize into discretional 'and temporary, "ad-hoc" network topologies, permitting individuals and device to scamlessly internetwork in areas with no pre-existing communication infrastructure, e.g., disaster recovery environments. And it's Associate in Nursinging autonomous system during which mobile hosts connected by wireless links area unit liberated to be dynamically and someday act as routers at together with network configuration could modification at any time, there's no direction or limit the movement so on, and therefore required a brand new facultative path agreement (routing protocol) to spot nodes for these actions communicate with one another path.

Keywords: Mobile Ad-hoc network, Routing protocol, Topologies, Bluetooth, MAC, Data security, Less power consumption, Local area network.

Introduction

The mobile Ad-hoc network is a collection of mobile devices which from a network with no pre-existing wiring or infrastructure. The system that enables wireless data communication is called network. The need for an application can give directions for the research enables new applications to be created. Mobile Ad-hoc network is a collection of two or more nodes or terminals with wireless communications and networking capability that communicate with each other without the aid of any centralized administrator also the wireless nodes that can dynamically form a network to exchange information without using any existing fixed network infrastructure. It consists of either computers, laptops, notebooks, routers, switches, cell phones, portable phones, PDA's, related operating systems / software, access points , base stations, antennas or towers etc. Various types of wireless networks being used now days are as; infrastructure-based WLAN, wireless Ad-hoc network, wireless personal area network, wireless cellular network, satellite system, television network and wireless sensor network (WSN) etc.

Basic research and potential applications of ad hoc networks are evolving together, spurring each other into further achievements. The need for an application can give directions for the research and, on

the other hand, the research enables new applications to be created. The dynamic source routing, Ad-hoc on-demand distance vector routing, and temporally ordered routing algorithm are some examples for the protocols that belong to this category. Recently, a new family of algorithms emerged inspired by swarm-intelligence, which provides a novel approach to distributed optimization problems. Ant based routing algorithms in ad-hoc networks, categorized to flat and hierarchical routing algorithms.

In the future, ad hoc networks probably form the outermost region of the internetwork, where a wired backbone connects both the fixed local area networks and the mobile both the fixed infrastructure and the ad hoc networks. Whereas the base stations of a fixed infrastructure networks are directly connected to the core, an Ad-hoc network is typically connected through a satellite link or a terrestrial switch fixed wired connection point, or mobile radio-link. MAC algorithm is similar to traffic regulations in the highway. Several vehicles cross the same road at a time but rules required to avoid collision e.g., follow the traffic lights, building the flyovers etc., Since ad hoc wireless network needs to address unique issues which are discuss in the next section, a different set of protocols is required for controlling access to the shared medium in such network and these are nothing but only MAC protocols. It is clear that ad hoc networking is in its infancy what comes to its research and further development are being expected for years to come. Hence, previous works considering the future research in this field have been very few in number. MAC protocols. The area is developing at the pace of several hundreds of publications yearly and the speed unavoidably outdates any summaries, conclusions and predictions in very short time. From the application point of view, another notable instance relevant the topic is the Wireless World Research Forum (WWRF) [6], a forum founded by several leading telecom corporations.

Routing Challenges in Ad-hoc Network

MANETs are internetworks formed by mobile wireless routers, with each router having one or more associated host devices. MANET's routing algorithms organize the network by automatically discovering the topology of the connectivity among constituent nodes.

The ability to independently self-organize and serve as its own infrastructure makes MANETs particularly attractive for the industrial communications requirements in harsh manufacturing environments. People develop lots of routing protocols to fit the mobility of the Wireless Ad-hoc Networks. The routing algorithms become more and more fit the rapid changing network topology of Wireless Ad-hoc Networks.

Self-organized networks can act in an independent way from any provider. Self-organized networks are also potentially very large and not regularly distributed. For example, one single network can cover the entire world. Also, self-organized networks are highly co-operative, the tasks at any layer are distributed over the nodes and any operation is the results of the cooperation of a group of nodes the most important characteristic which sets Wireless Ad-hoc networks apart from cellular networks is the fact that they do not rely on a fixed infrastructure. They also think Mobile Ad-hoc networks are very attractive for tactical communication in military and law enforcement.

Issues in AD-HOC Network

There are several issues within Ad-hoc networks that make them very complicated to integrate with the existing global internet. The problems are addressed below.

Routing

Routing is one of the most complicated problems to be solve as Ad-hoc networks have a seamless connectivity to other devices in its neighborhood. Because of multi hop routing no default route is

available. Every node acts a router and forwards each other's packets to enable information sharing between mobile nodes.

Security

Obviously a wireless link is much more vulnerable than a wired link. The science of cracking the encryption and eavesdropping on radio links has gone on since the first encryption of radio links was established. The user can insert spurious information into routing packets and cause routing loops, long time-outs and advertisements of false or old routing table updates. Security has several unsolved issues that are important to solve to make the ad hoc network into a good solution.

Quality of Service (QoS)

QoS is a difficult task for the developers, because the topology of an ad hoc network will constantly change. Reserving resources and sustaining a certain quality of service, while the network condition constantly changes, is very challenging. If the available capacity sets some limits for communications, so do the protocols as well. Route acquisition, service location and encryption key exchange is just examples of tasks that will require considerable overhead, which will grow rapidly with the network size. If the scarce resources are wasted with profuse control traffic, it is clear that ad hoc networks will see never dawn in practice. Scalability is an important research topic for the future, not only because of its necessity for ad hoc networks, but also because of the applicability of same ideas in the Internet. In the protocol design itself, several issues have to be considered with the potential applications in mind.

Protocol Techniques for MANET

Portable devices have limited capacity (battery power, available memory, and computing power) that further complicates the protocol design. Several protocols for ad hoc networks have been developed. The protocols can perform well under certain situations that they are designed to solve, but they fail completely in other situations that can occur in the network[2].

The routing protocols for ad hoc networks have been classified as:

Proactive/table driven

In proactive routing, each node has one or more tables that contain the latest information of the routes to any node in the network. Each row has the next hop for reaching to a node/subnet and the cost of this route. Various table-driven protocols differ in the way the information about change in topology is propagated through all nodes in the network. The two kinds of table updating in proactive protocols are the periodic update and the triggered update. Proactive routing tends to waste bandwidth and power in the network because of the need to broadcast the routing tables/updates. Furthermore, as the number of nodes in the MANET increases, the size of the table will increase; this can become a problem in and of itself. e.g. Destination Sequenced Distance Vector (DSDV) [3], Optimized Link State Routing(OLSR)[4].

Reactive/On-demand

They do not maintain or constantly update their route tables with the latest route topology. Instead, when a source node wants to transmit a message, it floods a query into the network to discover the route to the destination. The discovered route is maintained until the destination node becomes inaccessible or until the route is no longer desired. The protocols in this class differ in handling cache routes and in the way route discoveries and route replies are handled. Reactive protocols are generally considered efficient when the route discovery is employed rather infrequently in comparison to the

data transfer. Although the network topology changes dynamically, the network traffic caused by the route discovery step is low compared to the total communication bandwidth. e.g. Dynamic Source Routing Protocol (DSR) [5], Ad hoc On-Demand Distance Vector routing protocol (AODV), Temporally Ordered Routing Algorithm (TORA).

Current issues and proposed system

There are many open questions related to ad hoc networks applications. Before a public demand for any set of applications can be found, these networks will be deployed in various specialized cases. All proposed solution could be carefully tested and found out any constraint if any. In this section we discuss the issues that still remain a challenge for the researchers. The protocol dependent development possibilities are mostly omitted and the focus is on the major issues that stand in a way of having peer to peer connectivity everywhere in the future.

The topics are:

- Less Power consumption
- Data Security
- Node cooperation
- Client server model shift
- Quality of Service
- Interoperation with the internet
- Scalability
- Less Power consumption

Power Control

In routing, one usually tries to maximize the network lifetime. In other words, routes are selected by their transmission energy cost giving the priority to the nodes with full batteries. This way the time to network partition can be maximized distributedly. Further more, unicast and multicast routing should be considered separately when considering energy-efficiency due to the broadcast nature of the transmission [25].

Infrastructure Wireless Network

When a source node likes to communicate with a destination node, the former notifies the base station. At this point, the communicating nodes do not need to know anything about the route from one to another. All that matters is that the both the source and the destination nodes are within the transmission range for the Base Station and then if there's any one loses this condition, the communication will frustration or abort.



Figure 1 Infrastructure Wireless Network

AD-HOC Networking Research

As shown in the figure, the research activities will be grouped, according to a layered approach into three main areas:

- Enabling technologies;
- Networking;
- Middleware and applications.

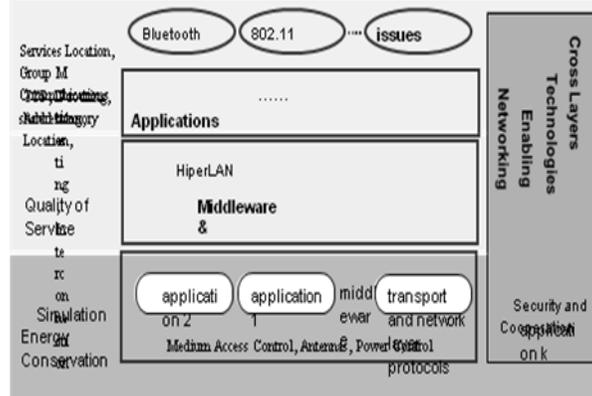


Figure 2 Ad-Hoc Network Implementation

Bluetooth

The Bluetooth technology is a de-facto standard for low-cost, short-range radio links between mobile PCs, mobile phones, and other portable devices [15,179]. The Bluetooth Special Interest Group (SIG) releases the Bluetooth specifications. Bluetooth specifications were established by the joint effort from over two thousand industry leading companies including Com, Ericsson, IBM, Intel, Lucent, Microsoft, Motorola, Nokia, Toshiba, etc.



Figure 3 Bluetooth Implementation

Conclusion

Whereas ad hoc networks will become widely used in military contexts in near future, the corporate world has to continue the daunting search for profitable commercial applications and possibilities of the technology. Meanwhile, the academic community has adopted the new field as a playground to apply their ideas to create something completely new. In all, although the widespread deployment of ad hoc networks is still years away, the research in this field will continue being very active and imaginative.

References

1. Imrich Chlamtac, Marco Conti, Jennifer J.-N. Liu, "Mobile Ad Hoc Networking: Imperatives And Challenges", Vol. 3, Issue 1, January -February 2003, Pp-13-65.
2. Saleh Ali K.Al-Omari¹, Putra Sumari², "An Overview Of Mobile Networks For The Existing Protocols And Applications", Vol. 2, Issue 1, March 2010, Pp-87-110.
3. Aleksi Penttinen, "RESEARCH ON AD HOC NETWORKING: CURRENT ACTIVITY AND FUTURE DIRECTIONS", Volume 4, Issue 5, December-2013,Pp-25-31.
4. Iqbaldeep Kaur, Navneet Kaur, Tanisha, Gurmeen, Deepi, "Challenges And Issues In Adhoc Network", Vol. 7, Iss Ue 4, Oct - Dec 2016, Pp-64-65.
5. Sumit Sharma, Ashutosh Dhamija And Vidhu Rawal, "Current Issues With Ad-Hoc Network & Proposed Solutions", Volume 7, Issue 10, December-2012, Pp-56-81.,
6. Mukesh Kumar , Rahul Rishi , D.K.Madan "Issues And Challenges Of Quality Of Service In Mobile Adhoc Network", Volume 1, Issue 3, December-2012, Pp-61-66.